



IMPERIAL AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.



14663

3-193-228-15-5000

Library,
Indian Agricultural Research Institute
New Delhi (India).

CHEMICAL REVIEWS

EDITORIAL BOARD

W. ALBERT NOYES, Jr., Editor

LOUISE KELLEY, Assistant Editor

WALTER G. WHITMAN

WENDELL M. LATIMER

V. S. CALCOTT

CARL S. MARVEL

HOWARD B. LEWIS

W. CONARD FERNELIUS

VOLUME 24

PUBLISHED BI-MONTHLY FOR
THE AMERICAN CHEMICAL SOCIETY
BY
THE WILLIAMS & WILKINS COMPANY
Baltimore, U. S. A.
1939

CONTENTS

NUMBER 1, FEBRUARY, 1939

| | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| The Fundamental Principles and Applications of Electrolysis with the Dropping Mercury Electrode and Heyrovský's Polarographic Method of Chemical Analysis. I M. KOLTHOFF AND JAMES J. LINGANE | 1 |
| Oxidation and Reduction of Organic Compounds at the Dropping Mercury Electrode and the Application of Heyrovský's Polarographic Method in Organic Chemistry. OTTO H. MULLER | 95 |
| The Electrodeposition of Hydrogen and Deuterium at the Dropping Mercury Cathode. J HEYROVSKÝ | 125 |
| The Chemistry of Retene DAVID E ADLSSON AND MARSTON TAYLOR BOGERT | 135 |

NUMBER 2, APRIL, 1939

| | |
|-----------------------------------------------------------------------------------------------------------------|-----|
| A Symposium on the Physical Chemistry of the Proteins | |
| Introduction to the Symposium. J W. WILLIAMS | 177 |
| Properties and Structure of Protein Monolayers. IRVING LANGMUIR AND VINCENT J SCHAFER | 181 |
| Some Physical-chemical Characteristics of Protein Molecules. EDWIN J. COHN | 203 |
| Theoretical Studies upon Dipolar Ions. J G. KIRKWOOD | 233 |
| The Application of the Theory of Absolute Reaction Rates to Proteins HENRY EYRING AND ALLEN E STEARN | 253 |
| Electrophoresis of Proteins by the Tiselius Method. L. G. LONGSWORTH AND D A. MACINNES | 271 |
| Opaque or Analytical Ultracentrifuges JAMES W. McBAIN | 289 |
| The Physical Chemistry of Tobacco Mosaic Virus Protein. MAX A. LAUFLER AND W. M. STANLEY | 303 |
| Chemical Aspects of the Precipitin and Agglutinin Reactions. MICHAEL HEIDELBERGER | 323 |
| The Polar Groups of Protein and Amino Acid Surfaces in Liquids. H. A. ABRAMSON, M H. GORIN, AND L S MOYER | 345 |

NUMBER 3, JUNE, 1939

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| The Wien Effect. Deviations of Electrolytic Solutions from Ohm's Law under High Field Strengths. HARTLEY C. ECKSTROM AND CHRISTOPH SCHMELZER | 367 |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-----|

Errata

Page 126. In the fourth line from the bottom of the page substitute “0.113” for “0.133”.

Page 128. The heading in the fifth column of table 1 should be ωi instead of i .

Page 129. In the second line from the bottom of the page substitute “more” for “greater.”

Page 131. The seventh equation should read

$$\pi - \pi' = RT/F \log \frac{a_1}{a_2} \cdot \frac{1 + \omega_2 i}{1 + \omega_1 i}$$

In the second line after this equation substitute “0.102” for “0.012.”

Page 134. Reference 4 should read “Trans. Faraday Soc. **33**, 660 (1937).”

THE FUNDAMENTAL PRINCIPLES AND APPLICATIONS OF ELECTROLYSIS WITH THE DROPPING MERCURY ELECTRODE AND HEYROVSKÝ'S POLAROGRAPHIC METHOD OF CHEMICAL ANALYSIS¹

I. M. KOLTHOFF AND JAMES J. LINGANE

*School of Chemistry, Institute of Technology, University of Minnesota,
Minneapolis, Minnesota*

Received December 18, 1938

CONTENTS

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------------|----|
| I. The principle of the polarographic method..... | 3 |
| II. The polarograph..... | 9 |
| III. Fabrication of the capillaries and the dropping electrode..... | 13 |
| IV. Types of polarographic cells..... | 15 |
| V. The potential of the quiet electrode and its significance in the evaluation of half-wave potentials..... | 17 |
| VI. The diffusion current..... | 19 |
| 1. The Ilković equation for the diffusion current | 20 |
| 2. Influence of the geometrical characteristics of the capillary on the diffusion current..... | 22 |
| 3. The product $m^{2/3} t^{1/6}$ as a function of the potential of the dropping electrode..... | 24 |
| 4. The relation between the diffusion current and the concentration of the reducible substance..... | 26 |
| 5. Direct test of the Ilković equation..... | 28 |
| 6. The "condenser current" and its influence in the measurement of very small diffusion currents..... | 31 |
| 7. The influence of temperature on the diffusion current..... | 36 |
| 8. Influence of the solvent on the diffusion current constant..... | 37 |
| VII. The migration current and its influence on the limiting current..... | 38 |
| 1. Relation between limiting currents in the absence of added salt and diffusion currents in the presence of an excess of indifferent salt... | 38 |
| 2. Increase, or "exaltation," of the migration current by the preceding discharge of an uncharged substance..... | 45 |
| 3. Exaltation of the migration current by the preceding discharge of a reducible ion | 48 |
| VIII. Maxima on current-voltage curves..... | 50 |
| 1. The Heyrovský-Ilković theory of maxima..... | 50 |

¹ Presented before the Division of Physical and Inorganic Chemistry at the Ninety-fifth Meeting of the American Chemical Society, held at Dallas, Texas, April 18-22, 1938.

| | |
|---------------------------------------------------------------------------------------------------|-----------|
| 2. Suppression and elimination of maxima..... | 56 |
| 3. Phenomena at the electrocapillary zero..... | 62 |
| 4. Influence of external resistance in the cell circuit on the maxima | 63 |
| IX. Equations of polarographic waves and the significance of the half-wave potential. | 67 |
| 1. Electrodeposition of simple metal ions..... | 67 |
| 2. Electrodeposition of complex metal ions..... | 75 |
| 3. Reduction of organic compounds | 78 |
| X. Anodic current-voltage curves with the dropping electrode..... | 79 |
| XI. Analytical applications of the polarographic method and polarometric titrations..... | 81 |
| 1. General applications..... | 81 |
| 2. Polarometric titrations..... | 87 |
| XII. Summary..... | 90 |
| XIII. References..... | 90 |

The polarographic method of chemical analysis, invented about sixteen years ago by Heyrovský (31), is based on the interpretation of current-voltage curves obtained by electrolyzing solutions of electroreducible or electrooxidizable substances, in a cell in which one electrode consists of mercury falling in small, slow drops from a glass tube with a very narrow capillary. From the current-voltage curves so obtained both the species and the concentration of the electroreducible or electrooxidizable substances in the solution can be determined, and under optimum conditions it is possible to detect and determine several substances simultaneously from a single c.v. (current-voltage) curve. The method is especially suitable for small concentrations (10^{-5} to 10^{-2} molar), and since the electrolysis can be carried out with a volume of solution as small as 0.1 cc. or even less, mere traces of substances can be determined. Since electrolysis is the most generally applicable method of carrying out oxidation-reduction reactions, it is evident that the number of substances subject to polarographic determination is potentially very great. The method can be used to detect and determine a large variety of organic substances, in addition to practically all the common electroreducible inorganic ions; hence it gives promise of becoming a universally applicable analytical technique in the various branches of chemistry and allied fields.

Although the polarographic method is a relatively new technique in electroanalytical chemistry, it has undergone a remarkably rapid development, due chiefly to the work of Heyrovský (38) and his school at the Charles University in Prague (Czechoslovakia) and of a few other investigators, notably Hohn (50) in Germany, Semerano (104) in Italy, and Shikata in Japan. Until recently the polarographic method was practically unknown in this country, and no complete review of the subject has heretofore been published. It is hoped that the present review of the fundamental principles of electrolysis with the dropping

mercury electrode, which is based largely on our own experience, will help to stimulate the interest of chemists in this versatile technique.

I. THE PRINCIPLE OF THE POLAROGRAPHIC METHOD

A typical arrangement for obtaining current-voltage curves with the dropping mercury electrode is shown schematically in figure 1. In this diagram A is the electrolysis cell containing the solution to be analyzed, B is the dropping mercury electrode, and C is a stationary pool of mercury on the bottom of the cell which acts as the second electrode. The dropping electrode consists of a drawn-out capillary tube, whose internal diameter at the tip is about 0.03 mm., connected to a reservoir of mercury. Mercury drops issue from the capillary at the rate of about one drop every

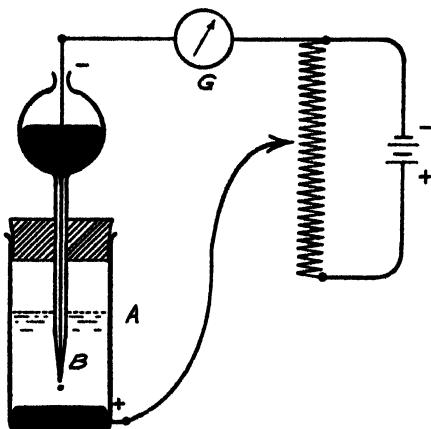


FIG. 1. Schematic representation of apparatus for obtaining current-voltage curves with the dropping electrode.

2 to 4 sec. The drops are very small, having a maximum diameter at the breaking point of only about 0.5 mm. The cell is connected, in series with a calibrated galvanometer G, to the battery and rheostat, by means of which an E.M.F. from zero up to the maximum E.M.F. of the battery can be applied to the cell. The applied E.M.F. is measured with a potentiometer, or, if the rheostat is a slide wire of uniform resistance, it can be computed in the usual way from the E.M.F. of the battery and the position of the sliding contact on the wire. The current-voltage curves are obtained by gradually increasing the applied E.M.F. and noting the current indicated by the galvanometer. The current is ordinarily quite small, seldom exceeding 50 microamperes (5×10^{-5} amperes).

The dropping electrode is usually connected to the negative pole of the

polarizing E.M.F., in which case cathodic current-voltage curves are obtained (electroreduction). In certain cases it is also possible to obtain anodic current-voltage curves (electrooxidation) by connecting the dropping electrode to the positive side of the polarizing E.M.F.

Since oxygen is readily reduced at the dropping electrode (124) and usually interferes with the current-voltage curves of other substances, it is generally necessary to remove dissolved air from the solution to be electrolyzed by bubbling an inert gas (nitrogen or hydrogen) through the cell before (but not during) the electrolysis.

The typical cathodic current-voltage curve shown in figure 2 was obtained by electrolyzing an air-free solution of 0.0013 M zinc sulfate in

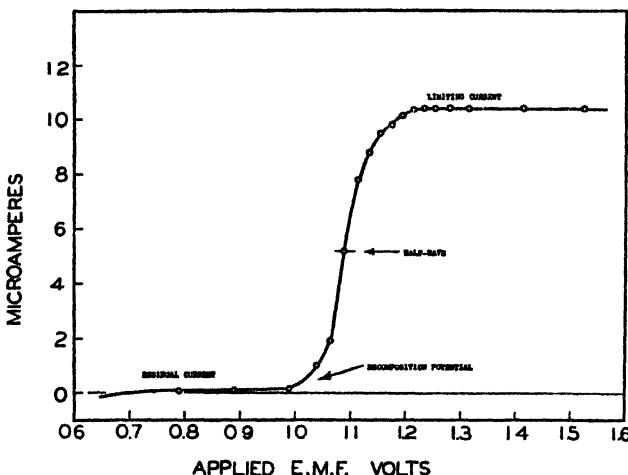


FIG. 2. Typical cathodic current-voltage curve obtained with the dropping mercury electrode.

0.1 N potassium chloride (65). In this polarogram the applied E.M.F. is plotted on the abscissa and the corresponding current on the ordinate. It will be noted that only an exceedingly small current, the "residual current," flowed through the cell until the decomposition potential was reached at an applied E.M.F. of about 1.0 volt. When the decomposition potential was exceeded continuous electrolysis began, involving the discharge of zinc ions at the dropping mercury cathode to form an extremely dilute zinc amalgam and the anodic dissolution of mercury, with the subsequent formation of calomel, at the large quiet anode. It will be noted that the current did not increase indefinitely with increasing applied E.M.F. after the decomposition potential was exceeded, but gradually approached a limiting value and finally became constant and independent of further increase in the applied E.M.F.

Under optimum conditions, and with all other factors constant, the limiting current is directly proportional to the concentration of the electro-reducible substance. This is the basis of quantitative polarography.

The limiting current is caused by a virtually complete state of concentration polarization at the dropping electrode. As a result of the discharge process, the concentration of the reducible material is depleted close to the surface of the dropping electrode, and this loss is compensated by diffusion of a fresh supply of reducible material from the body of the solution. The rate of diffusion depends directly on the difference in concentration between the depleted surface layer and the body of the solution. As the E.M.F. is increased above the decomposition potential, the current increases, the *average* concentration at the surface of the mercury drops decreases, and the rate of diffusion is correspondingly increased. As the applied E.M.F. is further increased, the concentration at the surface of the mercury drops becomes so small compared to the concentration in the bulk of the solution that the difference in concentration approaches a constant average value, equal simply to the concentration in the solution, and hence the rate of diffusion also becomes constant. From this point on the amount of material discharging, and thus the current, becomes constant and practically independent of further increase in the applied E.M.F. With an excess of some indifferent salt present in the solution, the limiting current is determined practically entirely by the rate of diffusion, and therefore is called a "diffusion current." Since the rate of diffusion attains a constant value, which is directly proportional to the concentration in the bulk of the solution, it is readily understandable why the diffusion current is directly proportional to the concentration in the body of the solution. A more complete quantitative interpretation of the diffusion current will be given in a later section.

With the dropping mercury electrode the current at each value of the applied E.M.F. is not constant but oscillates between a minimum and a maximum value as each mercury drop grows and falls. With a sensitive D'Arsonval galvanometer of relatively long period (ca. 15 sec.), the type usually employed to measure the current, the observed oscillations are not very large and the average current can be measured readily with a precision of better than 1 per cent.

It may be mentioned that the necessary conditions that must be fulfilled in order to obtain a limiting current are that at least one electrode of the cell be very small and that the concentration of the reducible substance be not too large. That is, conditions must be such as to favor the attainment of an extreme state of concentration polarization. Limiting currents are therefore not peculiar to the dropping mercury electrode, but may also be obtained with small solid electrodes, such as a platinum microelectrode, or with a small quiet mercury electrode. For example,

in figure 3 is given a cathodic current-voltage curve obtained by electrolyzing an air-free solution of $1.75 \times 10^{-3} M$ thallous chloride in 0.1 *N* potassium chloride, according to the technique already described, except that the dropping mercury electrode was replaced by a platinum microelectrode, consisting of a platinum wire 0.2 mm. in diameter sealed into the end of a glass tube in the usual way with 1 mm. left exposed (65). The first "wave" is due to the discharge of thallous ions and the second to the evolution of hydrogen. Mr. H. A. Laitinen in this laboratory has found that the diffusion current of thallous ions, and of various other ions, with a platinum microelectrode is directly proportional to the concentration of the ions.

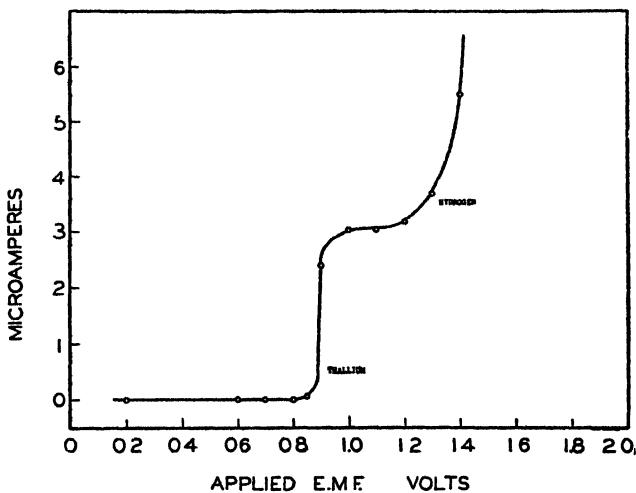


FIG. 3. Cathodic current voltage curve obtained with a platinum microelectrode

Although a solid microelectrode is in certain respects simpler to use than the dropping mercury electrode, it has several disadvantages. With a solid microelectrode the current does not become steady immediately at each new setting of the applied E.M.F., and several minutes are required to obtain a steady reading. Solid microelectrodes are also sensitive to vibrations, which disturb the diffusion and cause irregular fluctuations in the current. Furthermore, the deposition of base metals at a platinum microelectrode is hindered by the evolution of hydrogen. The dropping mercury electrode has none of these disadvantages. The (average) current becomes steady immediately at each new setting of the applied E.M.F. and is independent of the time of electrolysis. Owing to the high overvoltage of hydrogen on mercury, even the alkali and alkaline-

earth metals can be readily deposited, to give well-defined limiting currents (31, 89).

The favorable behavior of the dropping electrode is due to the fact that a fresh mercury surface is continuously being exposed by the growing drops; the phenomena at a given drop are exactly duplicated at its successor. Hence the current at any point on the current-voltage curve is solely a function of the potential of the dropping electrode, and is independent of the previous course of the electrolysis and of the length of time that the current has passed. A current-voltage curve obtained by increasing the E.M.F. in the usual way will generally be retraced exactly if the applied E.M.F. is gradually decreased to zero. These unique characteristics of the dropping electrode have made possible the development of the polarographic method.

The reason for using a relatively large, quiet pool of mercury as the second electrode of the cell is that such an electrode remains practically depolarized when the solution contains halide or other ions which form insoluble salts with mercury. It thus retains a practically constant potential independent of the applied E.M.F., and only the dropping electrode becomes polarized. For example, in solutions of chlorides the quiet electrode acquires a potential practically equal to the potential of a calomel electrode at the same chloride-ion activity.

The decomposition potential of a given solution is characteristic of the particular electroreducible substance present. Even more characteristic is the so-called "half-wave potential" (44), which, as its name implies, is the value of the applied E.M.F., or, better, the potential of the dropping electrode against an external reference electrode, at that point on a current-voltage curve where the current is equal to one-half of its limiting value (see figure 2). In contradistinction to the decomposition potential, which depends to some extent on the concentration of the reducible substance, the half-wave potential is, in general, entirely independent of the concentration of the reducible substance, provided that the composition of the solution with respect to foreign salts is kept constant (44). The half-wave potential of a given substance is also independent of the particular capillary used and of its drop time (44). The theory of the half-wave potential will be discussed in detail in a later section. It will suffice here to point out that qualitative polarographic analysis is based on the characterization of the half-wave potentials on the current-voltage curves of "unknown" solutions. By comparing the half-wave potential on an "unknown" current-voltage curve with the known values (38, 50, 44) for various substances, the substance responsible for the "polarographic wave" can be readily identified, and from the magnitude of the diffusion current its concentration can be determined.

When several electroreducible substances are present in the solution, each one will produce its own characteristic wave on the current-voltage curve, provided the half-wave potentials of the several substances are not too close together. The various half-wave potentials generally must differ from each other by at least 0.2 volt in order to obtain separate, well-defined waves. If the half-wave potentials of two substances are too close together, their waves will coalesce to give only a single wave on the current-voltage curve. In general, the half-wave potential of a given substance is not changed by the preceding discharge of another reducible substance. It may be noted, however, as one of the few exceptions to this general rule, that the half-wave potential of hydrogen discharge is decreased (shifted to more positive value) by adding to the

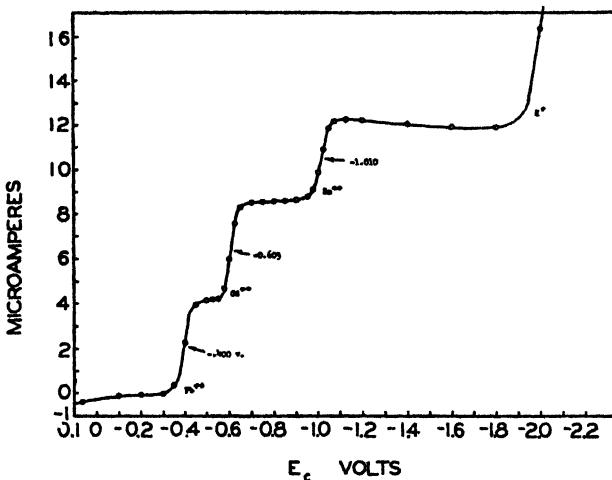


FIG. 4. Current-voltage curve of a solution containing several reducible ions

solution certain substances which decrease the overvoltage of hydrogen on mercury, e.g., traces of platinum metals (111, 112) and certain proteins and organic compounds containing sulfur (6).

The concentration of each substance in a mixture can be determined from its own wave height; thus it is often possible to obtain a simultaneous qualitative and quantitative analysis for several constituents in an unknown solution from a single current-voltage curve. For example, the current-voltage curve in figure 4 was obtained by electrolyzing an air-free 0.1 N potassium chloride solution which contained lead nitrate, cadmium sulfate, and zinc sulfate, each at a concentration of about 5×10^{-4} molar (65). The individual heights of the three waves are in each case proportional to the concentrations of the respective ions.

The half-wave potentials of many simple (hydrated) metal ions are shifted to more negative values when complex-forming salts are added to the solution. This fact is often used to separate the polarographic waves of two metal ions which would otherwise coincide. For example, the half-wave potential of simple lead ions is only about 50 millivolts more positive than that of simple thallous ions, and hence the waves of these two ions practically coincide. However, by adding a large amount of potassium tartrate or potassium cyanide to the solution, the half-wave potential of lead is shifted by several tenths of a volt to a more negative value, while that of thallium remains practically unchanged; therefore in such a solution two well-separated waves are obtained, which allow the analytical separation of the two metals. It is possible in many cases to determine the instability (complex) constant of the complex formed from the magnitude of the shift in the half-wave potential (44).

The current-voltage curves do not always have the "ideal" shapes shown in the preceding figures, but are often distorted by so-called "maxima." That is, the discharge curve rises sharply, but, instead of gradually developing into a limiting current, the current increases abnormally and practically linearly with the applied E.M.F. until a certain potential is reached, and then decreases more or less rapidly towards a limiting value. A more or less sharp maximum or peak in the current-voltage curve thus results (see section VIII). The characteristics of such maxima and methods of eliminating them, which are necessary for purposes of quantitative polarographic analysis, will be discussed in detail in a later section.

II. THE POLAROGRAPH²

In 1925 Heyrovský and Shikata (46) invented an instrument called the polarograph, which automatically obtains, and photographically records, current-voltage curves. Since Heyrovský and Shikata (46, 38) and Hohn (50) have given excellent detailed descriptions of the construction and manipulation of the polarograph, only the principle of the instrument need be discussed here.

The principle of the polarograph is shown schematically in figure 5.

² Various models of the Heyrovský polarograph may be purchased from the firm of Dre. V. and J. Nejedlý, Prague XIX, Vokovice, Husova tr. 76, Czechoslovakia. Polarographs are also manufactured by E. Leybolds Nachfolger A. G., Köln-Bayenthal, Bonner Str. 500, Berlin N. W. 7. The Leybold instrument has been described in detail by Hohn (50). The Leeds and Northrup Company of Philadelphia has constructed a few polarographs on special order, in which they incorporated the principles of their well-known "Micromax" potentiometer recorder. The details of this instrument may be obtained from Dr. V. M. Meloche of the University of Wisconsin or Dr. B. Clarke of the Bell Telephone Laboratories in New York City.

In this figure D is the electrolysis cell containing the solution to be analyzed. B is essentially a potentiometric bridge, consisting of a cylinder of insulating material wound with a rather low resistance wire (ca. 15 ohms) in exactly twenty turns, by means of which a variable E.M.F. is applied to the cell. Current for the bridge is obtained from the storage battery H, and the total potential drop across the bridge can be exactly adjusted to any desired value (usually 2 or 4 volts) by a regulating resistance and comparison against a Weston standard cell in the usual manner. Each turn of the bridge wire then corresponds to, let us say, 100 or 200 millivolts of applied E.M.F. The bridge is rotated by a small

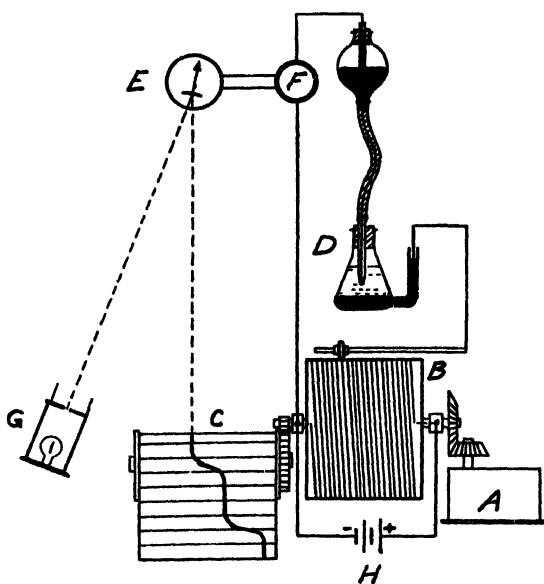


FIG. 5. Principle of the polarograph

electric motor A. A roll of sensitive photographic paper is carried by the cylinder C, enclosed in a light-tight housing, and connected by a system of gears to the bridge so that the two revolve simultaneously. The gear ratio is accurately adjusted so that one complete revolution of the bridge (100 or 200 millivolts) corresponds to about 1 cm. on the photographic paper. G is a powerful galvanometer light, the beam of which is reflected onto the photographic paper by the mirror of the sensitive D'Arsonval galvanometer E. The galvanometer is connected into the circuit in parallel with an Ayrton shunt F, by means of which its sensitivity may be regulated. When current flows through the galvanometer, the light beam traces a thin line on the photographic paper parallel to the axis of the roll.

With the cell in readiness and the dropping electrode in operation, the bridge is turned back until the moving contact (small grooved wheel) is at zero. The motor is then started, and the bridge and photographic roll revolve slowly at a uniform speed, which applies a uniformly increasing E.M.F. to the cell. As long as the applied E.M.F. is less than the decomposition potential, the current through the cell is very small (residual current), the beam of light from the galvanometer remains practically stationary, and the slow rotation of the drum traces a line along the circumference of the photographic roll. When the decomposition potential is reached, the increase in current through the cell circuit deflects the galvanometer, and the light beam traces a line approximately parallel to the axis of the roll; this is the discharge curve or polarographic wave. When the limiting current is reached and the light beam attains a constant displacement, the limiting current is traced as a circumferential line on the photographic roll.

The photographic roll is enclosed in a housing provided with a narrow collimating slit through which the light beam from the galvanometer enters. Each time the roll turns through a distance of 1 cm. (100 or 200 millivolts of applied E.M.F.) an auxiliary light, automatically flashed on, illuminates the entire length of the slit, and a thin line is printed on the paper. These lines mark the increments of applied E.M.F. The diffusion current is computed from the measured height of the polarographic wave and the known sensitivity of the galvanometer.

The polarograph has several advantages over manual measurement. In the first place, the curves can be obtained in a fraction of the time required to obtain them manually; a complete curve can be recorded in 10 min. or less. Several curves can be recorded on a single sheet of photographic paper, and a permanent record is thus obtained in a single operation. A continuous curve is obtained, instead of a series of points; hence very small waves, or slight peculiarities which might otherwise escape notice, are automatically detected.

Although a polarograph is very convenient, it is expensive and is not at all essential for obtaining current-voltage curves. For this reason we shall briefly describe a simple circuit which we have found very convenient for obtaining current voltage curves manually, and with which most of the current-voltage curves that we studied were obtained (65, 66).

The principle of the circuit is shown in figure 6. The E.M.F. applied to the cell was regulated by means of a rheostat, consisting of two radio rheostats of the "potentiometer" type, connected in series. A four-dial precision resistance box (9999 ohms), indicated by R in figure 6, was connected in series with the cell; by measuring the potential drop across this resistance with a potentiometer the current through the cell could be readily determined by applying Ohm's law. This method of measuring the current is very precise and accurate, but was somewhat unhandy.

To facilitate the current measurement a galvanometer, with suitable shunts for regulating its sensitivity, was also connected in series with the cell and the standard resistance. The E.M.F. applied to the cell, the potentials of the dropping electrode and of the quiet electrode against a saturated calomel reference electrode (S.C.E. in figure 6), and the potential drop across the standard resistance were all measured with a potentiometer. In order to facilitate these readings, a circular six-position double-pole

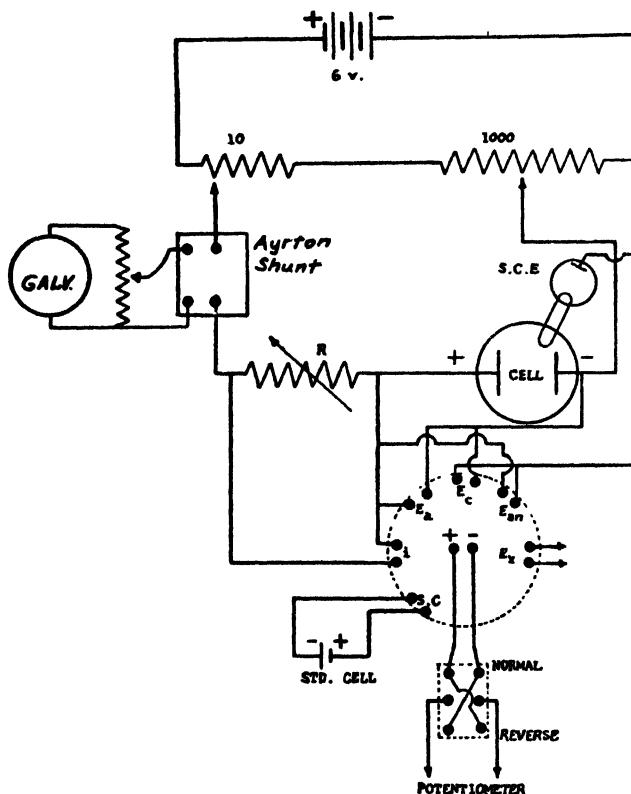


FIG. 6. Simple circuit for obtaining current-voltage curves manually

selector switch was used, as shown in figure 6. A Weston standard cell was also connected to one position of the selector switch for calibrating the potentiometer. Further details regarding the use of this simple circuit are described elsewhere (66).

Although the manual method of measurement was not as convenient as the use of the almost automatic polarograph, it possessed the advantage of somewhat greater precision and accuracy, which was more important

than convenience for our purpose of critically studying the polarographic method. The circuit in figure 6 has the additional advantage that the potentials of the dropping and quiet electrodes can be measured separately during the electrolysis. It should be mentioned that the magnitudes of the limiting and diffusion currents are entirely independent of the characteristics of the measuring circuit. On the other hand, the magnitude of the oscillations in the current during the formation of each mercury drop depends on the characteristics of the particular galvanometer used, and the heights of the maxima depend on the magnitude of the external resistance in the cell circuit (7), as will be discussed later.

Matheson and Nichols (78) have recently described a new type of polarograph, employing alternating current, and a cathode ray tube oscillograph. A similar instrument has also been described by Müller, Garman, Droz, and Petras (82). Schmidt (102) has recently described another modification of the Heyrovský polarograph.

III. FABRICATION OF THE CAPILLARIES AND THE DROPPING ELECTRODE

The capillaries are usually fabricated by drawing out the end of a 20-cm. length of Pyrex capillary tubing, of about 0.5 mm. internal diameter, until the internal diameter of the tip is about 0.03 to 0.04 mm. The end of the tube is thickened in the flame before drawing, so that the tip will not be too slender and fragile. The tip, which should have a uniform internal diameter, is cut off until the length of the narrowest part is about 2 cm. The exact dimensions must be decided by trial and error, until a capillary is obtained the drop time of which in 0.1 *N* potassium chloride can be readily adjusted to between 3 and 4 sec., by suitable adjustment of the pressure on the dropping mercury.

Maas (68) has made a very thorough study of the preparation of capillaries for the dropping electrode. He prepared the dropping electrode by cementing a 2- to 3-cm. length of commercial thermometer tubing (internal diameter about 0.03 mm.) into a wider glass tube, which was connected to the mercury reservoir. Siebert and Lange (109) also recommend this method; they used 6- to 8-cm. lengths of thermometer tubing of about 0.05 mm. internal diameter.

In order to obtain a strict linear relation between the diffusion current and the concentration of a given reducible substance, the drop time (time for the formation of each mercury drop) must not be too short. Maas (68) found that the drop time must be about 4 sec. or longer. We (65, 66) obtained a strict linear relation with capillaries the drop time of which in 0.1 *N* potassium chloride, with the dropping electrode disconnected from the polarizing E.M.F., was about 3 sec. If the drop time is too short, the diffusion current tends to be relatively too great at very small

concentrations, i.e., less than about $10^{-4} M$. On the other hand, if the drop time is too long, the current oscillations during the formation of each mercury drop are so great that it is difficult to measure accurately the average current. The optimum drop time is from 3 to 6 sec.

The capillary is usually connected to the mercury reservoir by a length of rubber pressure tubing, and the pressure on the dropping mercury can then be adjusted by simply raising or lowering the mercury reservoir. An ordinary leveling bulb is used as a mercury reservoir. This is a simple method, but it involves the danger of impurities from the rubber tube getting down into the very fine capillary tip and clogging it. In order to

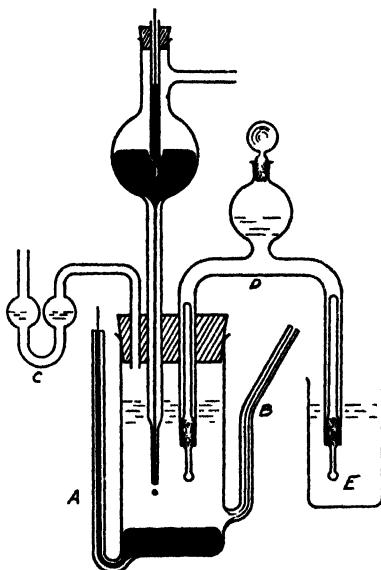


FIG. 7. Dropping electrode and electrolysis cell

avoid all possible contamination from rubber connections, we sealed the capillary directly to the mercury reservoir, as shown in figure 7. The pressure applied to the dropping mercury was regulated by means of two mercury leveling bulbs, as shown in figure 8. The total pressure on the dropping mercury is equal to the height of the mercury column in the dropping electrode itself, plus or minus the pressure indicated by the open-end mercury manometer. The pressure required to operate the dropping electrode depends, of course, on the characteristics of the particular capillary used; it is usually between 20 and 60 cm. of mercury.

When not in use the tip of the dropping electrode should be kept immersed in pure mercury, with the pressure reduced to prevent the mercury

in the reservoir from draining out. Before the electrode is placed in a solution, the pressure is increased so that the mercury drops issue from the capillary and thus prevent the solution from entering the tip. After long use the capillary may become partially dirty on the inside and behave erratically. When this happens, the mercury is removed and aqua regia is sucked through the tip into the bulb, by partially evacuating the latter. The aqua regia is then removed by thorough rinsing with pure water, and the capillary is dried by drawing a current of dry *filtered* air through it. With these simple precautions it is possible to use a single capillary indefinitely. At the time of this writing we are using a capillary that has been in daily use for over a year, during which time it has been cleaned several times without noticeably changing its properties.

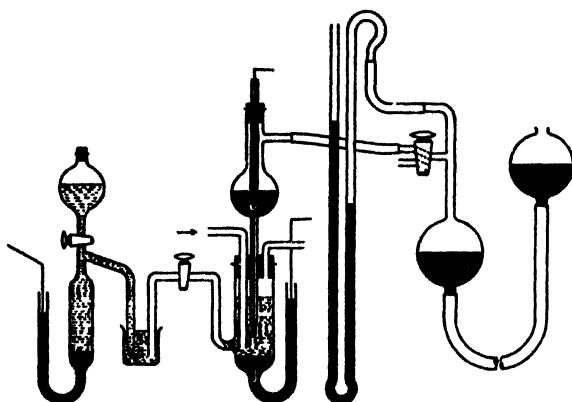


FIG. 8. Arrangement for regulating the pressure on the dropping mercury

IV. TYPES OF POLAROGRAPHIC CELLS

A few of the most common types of polarographic cells that have been described in the literature (mainly by Heyrovský) are shown in figure 9, which has been taken from the monograph of Hohn (50). These cells are very convenient for practical analytical work. Cells d and e are of the type most commonly used; electrical connection to the quiet electrode is made by means of a platinum wire sealed through the bottom of the cell as shown, and the side tubes are for the ingress and exit of nitrogen or hydrogen. Cells c and g are for the electrolysis of a very small volume of solution, and the arrangement shown under b is for the electrolysis of a single large drop. An ordinary small beaker may serve as a simple cell when the electrolysis can be carried out without removing air from the solution, as shown under a. Cell f is used for collecting and weighing the mercury drops. The vessel h is for freeing the mercury for the quiet

electrode from air before introducing it into the cell; this is often desirable. The vessel, containing the proper amount of mercury, is connected into the gas line with one side tube connected to the gas inlet tube of the cell. After nitrogen or hydrogen has been passed through the system, the vessel is inverted so that the de-aired mercury flows into the cell.

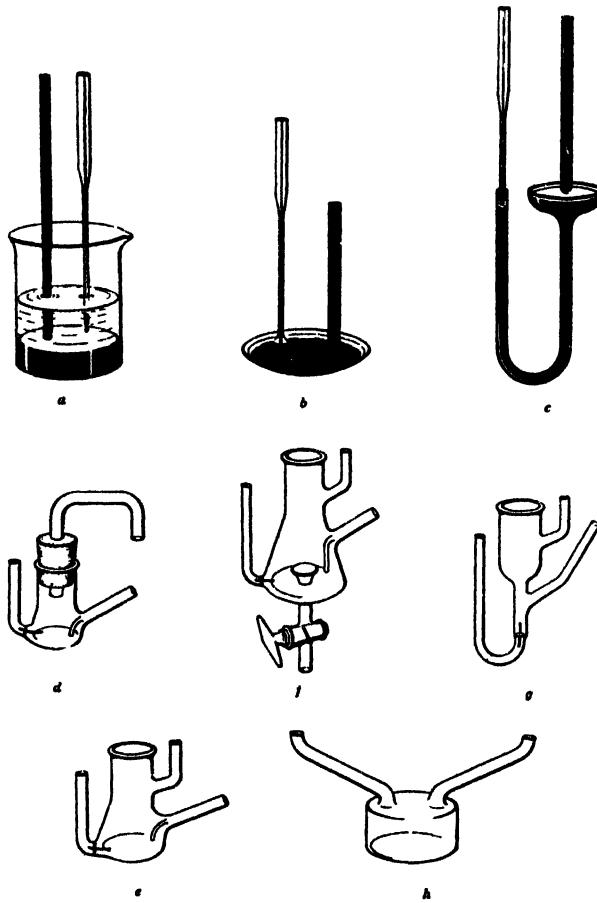


FIG. 9. Various types of polarographic cells

Majer (74) has recently described in detail various types of cells that are particularly suited for microchemical analysis, as well as other special apparatus for micropolarographic work.

In certain cases it may be necessary to work with the cell in a thermostat and to measure the potentials of the dropping electrode and of the quiet electrode against an external calomel reference electrode during the

electrolysis. The cell shown in figure 7 (65, 66) is very suitable for these purposes. Mercury for the quiet electrode is poured into the cell until it rises part way into the capillary side tube A, and electrical connection is made by means of a platinum wire immersed in the mercury in this side tube. This method of making electrical connection eliminates troublesome platinum-Pyrex seals. The capillary side tube B is for the introduction of nitrogen or hydrogen, and C is a gas outlet trap, containing either pure water or a little of the cell solution. The salt bridge D, designed by Irving and Smith (58), is usually filled with the same solution as in the cell. The ground-glass plugs in the ends of this bridge effectively prevent mixing of the bridge and cell solutions. The external end of the bridge dips into an intermediate vessel E, filled with concentrated potassium chloride solution, into which the side arm of the calomel electrode also dips.

As already mentioned, oxygen is reduced at the dropping electrode; hence dissolved air must usually be removed from the solution to prevent the oxygen wave from masking the waves of other substances. The time required to remove dissolved air, by bubbling nitrogen or hydrogen through the solution, is usually about 15 to 30 min. The gas stream is stopped before the electrolysis, because its stirring effect causes irregular fluctuations in the current.

Instead of using a quiet mercury electrode as the second electrode of the cell, it is often more convenient to use an external calomel, or mercury-mercurous sulfate, electrode for this purpose, as recommended by Hohi (50) and Maassen (69). The external electrode is connected to the cell by a low-resistance salt bridge. This method is especially suitable for practical routine work, because it eliminates the use of the extra mercury for the quiet electrode.

V. THE POTENTIAL OF THE QUIET ELECTRODE AND ITS SIGNIFICANCE IN THE EVALUATION OF HALF-WAVE POTENTIALS

The half-wave potentials of a current voltage curve are characteristic of the particular electroreducible, or electrooxidizable, substances present, provided the values are referred to an external reference electrode of constant known potential. However, when the current-voltage curves are obtained by use of the polarograph, the observed half-wave values are in terms of the total applied E.M.F. rather than of the potential of the dropping electrode itself; hence they depend on the potential of the quiet electrode as well as on that of the dropping electrode. Since the potential of the quiet electrode is a variable quantity, depending on the nature and concentration of the foreign salts present in the solution, the values of the half-wave applied E.M.F. are not characteristic of the reducible sub-

stances present. However, if the potential of the quiet electrode is known, the characteristic half-wave potentials of the dropping electrode can be deduced from the half-wave values of the total applied E.M.F.

If the potential of the dropping electrode is measured against an external reference electrode during the electrolysis (manual measurement), or if the external electrode is actually employed as the second electrode of the cell (measurement with the polarograph), the characteristic half-wave potentials will, of course, be obtained directly. This is the method that we prefer, but it is not common polarographic practice.

The usual practice is to measure the potential of the quiet electrode (usually the anode) against an external calomel reference electrode, either at the beginning or at the end of an experiment, and the potential of the dropping electrode is then computed from the simple relation,

$$E_a = E_{q.e.} - E_{d.e.} \quad (1)$$

In this equation E_a is the total E.M.F. applied to the cell, $E_{d.e.}$ is the potential of the dropping electrode, and $E_{q.e.}$ is the potential of the quiet electrode. The signs of $E_{d.e.}$ and $E_{q.e.}$ are taken to be negative when the electrode in question is negative to the calomel reference electrode, and positive when the electrode is positive to the calomel electrode. It is usually assumed that the potential of the quiet electrode remains constant during the electrolysis, and the half-wave potentials of the dropping electrode are then computed from the half-wave values of the applied E.M.F. by means of equation 1.

Majer (76) has shown that the quiet electrode remains depolarized during the electrolysis (retains a constant potential) if its area is greater than about 1 cm^2 , and also if the solution contains halide ions or other ions that form insoluble salts with mercury. Our experience is in accord with that of Majer.

When the mercury for the quiet electrode is added to a dilute chloride solution, sufficient mercurous chloride is formed by interaction of the mercury with chloride ions, hydrogen ions, and dissolved oxygen to saturate the solution with calomel before the air is removed, and the quiet electrode thus acquires the potential of a calomel electrode corresponding to the particular chloride-ion activity. The same principle applies when the solution contains iodide, bromide, hydroxide, and various other ions that form insoluble compounds with mercury. Since the current passing through the polarographic cell is very small, usually less than 50 microamperes, the current density at the relatively large quiet electrode is exceedingly small. In a solution already saturated with a mercurous salt, the only effect of this small current at the quiet electrode will be to increase or decrease the amount of the solid mercurous salt, with no

appreciable change in the potential-determining activity of the mercurous ion, and therefore the potential of the quiet electrode remains constant during the electrolysis.

VI. THE DIFFUSION CURRENT

It has already been mentioned that the phenomenon of a limiting current is caused by an extreme state of concentration polarization at the dropping electrode. When a limiting current is reached, the average concentration of the electroreducible substance at the surface of the mercury drops has been reduced to a very small minimum value, which is negligibly small compared to the concentration in the bulk of the solution. In other words, at potentials corresponding to the limiting current, the reducible substance is being reduced as fast as it reaches the surface of the dropping electrode. Under these conditions the current will evidently be determined by the rate of supply of the reducible substance from the body of the solution.

In the most general case reducible ions will be drawn to the depleted region very close to the surface of the dropping electrode by two forces: (1) a diffusive force, proportional to the concentration gradient between the surface of the mercury drops and the body of the solution, and (2) an electrical force, proportional to the electrical potential gradient around the mercury drops. In general, then, reducible ions are supplied to the surface of the dropping electrode partly by diffusion and partly by electrical migration, so that the limiting current can be regarded as the sum of a "diffusion current" and a "migration current" (53).

The current through an electrolyte solution is carried impartially by all the ions present. The fraction of the total current carried by any particular species of ion depends on its relative concentration in the solution, and, to a lesser degree, on its valence and intrinsic mobility; in other words, on its transference number in the particular solution in question. Therefore, if a large excess of an indifferent³ salt is added to a solution of a reducible salt, the current through the solution will be carried practically entirely by the ions of the added salt. Under these conditions electrical migration of the reducible ions, and hence the migration component of the limiting current, will be practically completely eliminated,

³ An indifferent salt, in electroreductions, is one that is reduced at a more negative potential than the reducible substance in question, and one that does not form complexes, or react specifically in other ways, with the reducible substance. Because of their very negative reduction potentials, the salts of the alkalies and alkaline earths and ammonium salts are usually used as indifferent salts. Tetrasubstituted alkyl-ammonium salts are commonly used in polarographic analysis for the alkali and alkaline-earth ions.

and the limiting current will be solely a diffusion current. From an analytical viewpoint the most important characteristic of the diffusion current is that, with all other factors constant, it is directly proportional to the concentration of the reducible ions.

The diffusion current, at a given concentration, is influenced by various factors, such as the geometrical properties of the particular capillary used, the pressure on the dropping mercury and the drop time, the diffusion coefficient of the reducible substance, and the temperature. In this section the influence of these factors will be discussed in some detail. The characteristics of the migration current will be considered later.

It may be well to emphasize that the current measured with the dropping electrode is an average value during the life of the mercury drops, that is, it is the average of the observed current oscillations as each mercury drop is formed. From the investigations of Ilković (54), who measured the actual change in current during the life of individual drops with an oscillographic arrangement, and from the theoretical discussion given below, it is known that the current often increases many fold during the life of each mercury drop. The actual variation depends on many factors, such as the concentration of both the reducible and foreign substances, the potential of the dropping electrode, and the resistance in the cell circuit (7, 54). However, the current oscillations occur with a period corresponding to the drop time, which is usually so much shorter than the period of the current-measuring galvanometer that the latter cannot follow the entire change in current, but simply oscillates with relatively small amplitude around the average value.

1. The Ilković equation for the diffusion current

Ilković (53) has derived an equation for the diffusion current in terms of the concentration of the reducible substance, its characteristic diffusion coefficient, the amount of mercury flowing from the capillary per second, and the drop time. The Ilković equation has recently been re-derived in a more exact way by MacGillavry and Rideal (72). Although space does not permit the detailed reproduction of these derivations, we shall outline the essential steps involved.

When an excess of some indifferent salt is present in the solution, the diffusion current will be determined by the amount of the reducible substance that diffuses up to the surface of the dropping electrode per second. The instantaneous value of the diffusion current, at any instant during the life of a mercury drop, is given by

$$i = nF(4\pi r^2f) \quad (2)$$

in which n is the number of electrons involved in the reduction of one molecule of the reducible substance (or the number of faradays of elec-

tricity required for the reduction of one mole), F is the faraday, r is the radius of the mercury drop at any time t^* , during its life, measured from the beginning of the drop formation, and f is the flux of the reducible substance (amount diffusing per square centimeter per second) *at the mercury-solution interface* (53). The method of derivation used by Ilković and by MacGillavry and Rideal consists in obtaining expressions for the flux, f , and the radius of the drop, r , from which an instantaneous value of the diffusion current can be computed. The average value of the diffusion current is then obtained by integrating over the life of the drop.

Because the expansion of the growing mercury drops causes a corresponding expansion of the surrounding solution, the diffusion of the reducible substance takes place in a medium that is moving with respect to the surface of the mercury drops. The diffusion layer around each mercury drop is a spherical shell, which expands and becomes thinner as the drop grows. The concentration of the reducible substance varies from a relatively very small value at the mercury-solution interface to its concentration in the body of the solution at the outside of the diffusion shell. On the basis of these considerations Ilković and MacGillavry and Rideal derived the following expression for the instantaneous current at any time, t^* , during the life of a drop.

$$i = 4 \left(\frac{7\pi}{3} \right)^{1/2} nFD^{1/2}C \frac{r^2}{(t^*)^{1/2}} \quad (3)$$

In this equation C is the concentration of the reducible substance in the body of the solution, r is the radius of the mercury drop at the time t^* , and D is the diffusion coefficient of the reducible substance. The *average* current during the life of each drop, designated by i_d , is given by

$$i_d = (1/t) \int_0^t i dt = 24\pi^{1/2} nFD^{1/2}Cr_{\max}^2/(21)^{1/2} t^{1/2} \quad (4)$$

in which t is the drop time and r_{\max} is the maximum radius of the drops.

If m is the weight of mercury flowing from the capillary per second, then the weight of an individual mercury drop is mt , and the volume of each drop, assumed to be spherical, will be given by

$$V = (4/3)\pi r_{\max}^3 = mt/d \quad (5)$$

where d is the density of mercury (13.6 g. per cubic centimeter). Hence, in terms of m and t , the maximum radius of each drop is

$$r_{\max} = (3mt/4\pi d)^{1/3} \quad (6)$$

By substituting this expression for r_{\max} into equation 4 and collecting all the numerical constants into a single term, we finally obtain

$$i_d = 0.627nFD^{1/2}Cm^{2/3}t^{1/6} \quad (7)$$

which is the equation for the diffusion current derived originally by Ilkovič. As will be shown later, diffusion currents calculated from the Ilkovič equation are, in most cases, in satisfactory agreement with the experimentally observed values.

We shall now consider the influence of the various factors, n , D , C , m , and t , on the diffusion current.

2. Influence of the geometrical characteristics of the capillary on the diffusion current

The quantities m and t in the Ilkovič equation obviously depend on the geometrical characteristics of the particular capillary used; they also depend on the pressure on the dropping mercury and, to a lesser degree, on the temperature. The observed effects can be interpreted satisfactorily by making use of the familiar hydrodynamical equation of Poiseuille,

$$W = \pi d P R^4 t / 8 L \eta \quad (8)$$

In this equation W is the weight of a liquid of density d and viscosity coefficient η , that flows in t seconds from a capillary tube of radius R and length L under the pressure P . In the case of the dropping electrode, if t is taken as the drop time, then W will be the weight of a single drop, and the weight of mercury flowing from the capillary per second will be given by

$$m = W/t = \pi d P R^4 / 8 L \eta \quad (9)$$

According to this equation m will be directly proportional to P , and t will be inversely proportional to P . In view of these relations the product $m^{2/3} t^{1/6}$ becomes

$$m^{2/3} t^{1/6} = (k' P)^{2/3} (k''/P)^{1/6} = k P^{1/2} \quad (10)$$

and the Ilkovič equation becomes

$$i_d = 0.627 k n F D^{1/2} C P^{1/2} \quad (11)$$

Therefore, if all other conditions are kept constant, the diffusion current should be directly proportional to the square root of the pressure on the dropping mercury. Ilkovič (53) and Maas (68), particularly the latter, have shown that this is actually the case.

For example, Ilkovič (53) found the following values of the diffusion current of 0.001 M mercuric chloride in 0.01 N sodium chloride, at various values of the pressure, P , on the dropping mercury.

| | | | | | |
|---------------|------|------|------|------|------|
| i_d in mm. | 28.0 | 30.5 | 33.0 | 35.5 | 37.5 |
| P in cm. | 39.1 | 46.3 | 53.2 | 62.4 | 70.1 |
| $i_d/P^{1/2}$ | 4.48 | 4.49 | 4.46 | 4.49 | 4.48 |

The diffusion current is expressed in terms of millimeters of galvanometer deflection. As predicted by equation 11, the ratio $i_d/P^{1/2}$ was constant over a range of pressures that varied almost twofold.

From an analytical viewpoint an important conclusion to be drawn from the foregoing relations is that two different capillaries will give the same diffusion current in a given solution when the product $m^{2/3}t^{1/6}$ is the same for both. This conclusion has been verified with many different capillaries by Maas (68). For example, he obtained the data in table 1 with six different capillaries. These data were obtained by electrolyzing 0.005 N solutions of cadmium chloride in 0.1 N potassium chloride (air-free) at 25°C. It is seen that the linear relation between i_d and $m^{2/3}t^{1/6}$ is obeyed with an accuracy of about ± 1.5 per cent.

TABLE 1
Linear relation between i_d and $m^{2/3}t^{1/6}$ (Maas)

| CAPILLARY | <i>t</i> | $m^{2/3}t^{1/6}$ | i_d | $i_d/m^{2/3}t^{1/6}$ |
|--------------|----------|------------------|-------|----------------------|
| 1 | 6.43 | 0.784 | 6.92 | 8.83 |
| 2 | 9.28 | 0.814 | 7.26 | 8.92 |
| 3 | 5.90 | 0.854 | 7.88 | 9.23 |
| 4 | 4.89 | 1.108 | 10.05 | 9.07 |
| 5 | 4.10 | 1.288 | 11.73 | 9.11 |
| 6 | 3.17 | 1.508 | 13.82 | 9.16 |
| Average..... | | | | 9.04 ± 0 |

It should be noted that the foregoing values of t and i_d were measured at a potential of the dropping electrode of -0.75 volt against a normal calomel electrode. As will be shown later, t , and hence the product $m^{2/3}t^{1/6}$, depend on the potential of the dropping electrode; therefore these quantities should be determined at the potential at which the diffusion current is measured.

It is interesting to compare the foregoing data of Maas with corresponding data obtained by Lingane (65) under the same conditions. Lingane electrolyzed a 0.005 N (0.0025 M) solution of cadmium sulfate in 0.1 N potassium chloride at 25°C., and found a diffusion current of 20.0 microamperes, with a capillary for which m was 2.626 mg. sec.⁻¹ and t was 2.88 sec. at $E_{d..} = -0.8$ volt against a saturated calomel electrode. Hence, in Lingane's experiment, the product $m^{2/3}t^{1/6}$ was equal to 2.28 mg.^{2/3} sec.^{-1/2}, and the quotient $i_d/m^{2/3}t^{1/6}$ was $20/2.28 = 8.77$. This quotient agrees to within about 3 per cent with the foregoing average value found by Maas, even though the drop time was much shorter than,

and the product $m^{2/3}t^{1/6}$ was about twice as great as, the average values of these quantities in Maas' experiments. This comparison of entirely independent data of different workers constitutes further confirmation of the linear relation between i_d and $m^{2/3}t^{1/6}$ predicted by the Ilkovič equation.

We may therefore state as a fairly accurate rule (within ca. ± 3 per cent) that *diffusion currents obtained under otherwise identical conditions with various different capillaries will bear the same ratios to each other as the products $m^{2/3}t^{1/6}$ of the various capillaries, provided the drop time is at least 3 sec. or greater.* The importance of this rule in quantitative polarography is at once evident, since it makes possible a simple correlation of diffusion currents obtained with different capillaries.

It is strongly recommended that future workers in the field of polarography include data for m , t , and the product $m^{2/3}t^{1/6}$ in their reports, so that the diffusion current data will be on a common basis for comparison. It has been more or less common practice to report diffusion current data in terms of millimeters of galvanometer deflection, and in many cases the sensitivity of the galvanometer has not been specified. It is hoped that future authors will convert such data to microamperes in their published reports, so that their results can be compared with those of other workers.

3. The product $m^{2/3}t^{1/6}$ as a function of the potential of the dropping electrode

It has been shown by various workers, Ilkovič (53), Maas (68), Kucera (63), and Lingane (65, 66), among others, that at a given pressure the amount of mercury flowing from a capillary per second, m , is practically constant and is independent of the potential of the dropping electrode. On the other hand, the weight of each mercury drop, W , and the drop time, t , are practically directly proportional to the interfacial tension at the mercury-solution interface (63). As is well known, the interfacial tension, σ , at a mercury-solution interface, and therefore t , vary with the potential of the dropping electrode; the parabolic curve obtained by plotting σ or t against the potential of the dropping electrode is the so-called electrocapillary curve of mercury. It is to be expected, therefore, that $m^{2/3}t^{1/6}$ will vary with the potential of the dropping electrode, approximately according to the sixth root of σ , and that i_d will vary correspondingly and not be strictly constant at various values of E_d .

The extent to which m , t , and the product $m^{2/3}t^{1/6}$ vary with the potential of the dropping electrode, in air-free 0.1 N potassium chloride solution, is illustrated by the data in table 2 (65, 66; see also reference 53).

The collection of a specified number of mercury drops, required for the determination of m , was accomplished without admitting air to the cell and without removing the cell from the thermostat, by means of the

simple arrangement shown in figure 10. The detailed manipulations are described elsewhere (66).

It will be noted from the data in table 2, that m was practically constant and independent of the potential of the dropping electrode. The drop

TABLE 2

The product $m^{2/3}t^{1/6}$ as a function of the potential of the dropping electrode in 0.1 N potassium chloride

$P = 21.8 \pm 0.2$ cm.; temperature = 25°C.; air was removed from the solution with purified nitrogen

| $E_{d.e.}$ volts vs. S.C.E. | i microamperes | m mg. sec. ⁻¹ | t seconds | $m^{2/3}t^{1/6}$ mg. ^{2/3} sec. ^{-1/2} |
|--------------------------------|---------------------|-------------------------------|----------------|-------------------------------------------------------------|
| 0.0 | -0.18 | 2.651 | 2.75 | 2.27 |
| -0.5 | +0.10 | 2.621 | 3.08 | 2.29 |
| -0.8 | 0.21 | 2.626 | 2.88 | 2.28 |
| -1.1 | 0.29 | 2.625 | 2.66 | 2.24 |
| -1.5 | 0.38 | 2.672 | 2.18 | 2.18 |
| -1.9 | 0.70 | 2.681 | 1.56 | 2.08 |

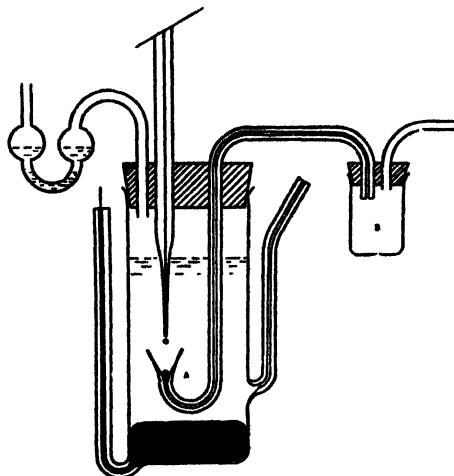


FIG. 10. Arrangement used to collect mercury drops for the determination of the amount of mercury flowing from the capillary per second.

time, t , on the other hand, first increased, then passed through a maximum, and finally decreased quite rapidly, with increasing negative potential. The drop weight behaved similarly. A plot of either the drop weight, or t , against $E_{d.e.}$ gives the electrocapillary curve in this solution.

The change in the product $m^{2/3}t^{1/6}$ with increasing negative potential

followed the same general course as the electrocapillary curve, but the variation in $m^{2/3}t^{1/6}$ was, of course, much less, because it depends on only the sixth root of t . At potentials from zero up to about -1.0 volt, the product $m^{2/3}t^{1/6}$ was constant to within about ± 1 per cent, but at potentials more negative than -1.0 volt it decreased rapidly.

It was also found (65, 66) that the product $m^{2/3}t^{1/6}$ was the same in 0.001 N as in 1 N chloride solutions, and also the same as in 0.1 N potassium chloride at the same potential when cadmium or thallous ions were discharging. Furthermore, the presence of oxygen and its simultaneous discharge with thallous ions had no appreciable effect on the product $m^{2/3}t^{1/6}$.

In addition to being independent of the potential of the dropping electrode, m is also practically independent of the medium in which the mercury drops form, that is, independent of the interfacial tension at the surface of the mercury drops, provided the pressure and the temperature are kept constant. This is demonstrated by the following data, which we obtained when the mercury drops were forming in various media. The same capillary was used as in table 2, and the pressure was kept constant at 21.8 cm.

| MEDIUM | t seconds | m $mg. sec.^{-1}$ |
|---------------------|----------------|------------------------|
| | | |
| Water | 3.20 | 2.69 |
| 0.5 N KCl | 3.00 | 2.65 |
| 0.1 N KI | 2.39 | 2.64 |
| Air | 20.1 | 2.64 |

According to the Ilković equation i_d is directly proportional to $t^{1/6}$, and therefore it is to be expected that i_d will not be strictly constant at various potentials, but should decrease appreciably at potentials more negative than about -1.0 volt. This effect has actually been demonstrated by Ilković (53). For example, in the electrolysis of 0.001 M mercuric chloride in 0.01 N sodium chloride, he obtained the data given in table 3. It is evident that the decrease in i_d with increasing negative potential must be taken into account when comparing the diffusion currents of the same concentration of two different substances, if the half-wave potentials of the two substances are far apart.

4. The relation between the diffusion current and the concentration of the reducible substance

With all other conditions constant, it is evident from equation 7 that the diffusion current should be directly proportional to the concentration

of the reducible substance. This is, of course, the most important characteristic of the diffusion current from an analytical viewpoint. The linear relation between i_d and C has been verified by many workers, in the reduction of both organic substances and inorganic ions (38, 50, 65, 66, 68).

We (65, 66) tested the relation between i_d and C in the discharge of thallous, lead, cadmium, zinc, iodate, and ferricyanide ions from 0.1 N potassium chloride solutions containing a trace of methyl red to eliminate the maxima. In the case of the four metal ions the linear relation was obeyed within about ± 1.5 per cent over a concentration range from 10^{-4} to 10^{-2} molar, after proper correction was made for the residual current found with the 0.1 N potassium chloride solution alone (table 2). In the case of iodate ions the linear relation was obeyed from 10^{-4} to 10^{-3} molar, and in the case of ferricyanide ions from 10^{-4} to 3×10^{-3} molar (65, 66).

TABLE 3
Variation of i_d with the potential

| <i>E_{d.o.}</i> | <i>t</i> (TEN DROPS) | <i>i_d</i> | <i>i_d/t^{1/2}</i> |
|-------------------------|----------------------|----------------------|--------------------------------------|
| volts | seconds | mm. | |
| -0.85 | 20.3 | 31.2 | 18.89 |
| -0.85 | 19.7 | 31.0 | 18.88 |
| -1.05 | 18.55 | 30.8 | 18.93 |
| -1.25 | 17.5 | 30.4 | 18.87 |
| -1.45 | 15.7 | 30.0 | 18.96 |
| -1.65 | 13.65 | 29.5 | 19.06 |

In order to obtain strict proportionality between i_d and C , the drop time should be at least 3 sec. or longer (66, 68). The linear relation seems to hold for all cases in which the electroreduction is reversible. When the electroreduction is irreversible, as it is in the case of nitrate ions, iodate ions, and many organic compounds, the linear proportionality may hold only under certain specified conditions, and only over a limited range of concentrations.

Certain investigators, notably Hamamoto (20) and Maas (68), found that the diffusion currents at very small concentrations (less than about 10^{-4} molar) were greater than would correspond to strict proportionality with the concentration. When the concentration of the reducible ions is very small, the diffusion current is usually not strictly constant but often increases with increasing negative potential. Several empirical methods for evaluating the diffusion current under such unfavorable conditions have been discussed by Hohn (50), Maas (68), and Borcherdt,

Meloche, and Adkins (5). It should be realized that the apparent diffusion current is the sum of the true diffusion current and a "residual current." In the absence of electroreducible substances (e.g., in the electrolysis of a pure solution of potassium chloride), the residual current is due to the so-called "charging current," or "condenser current," which is discussed in a later section. The residual current increases with increasing negative potential, and therefore, even though the true diffusion current may remain constant, the apparent diffusion current tends to increase with increasing negative potential. The best method for evaluating the true diffusion current is to determine the residual current in the medium in which the electrolysis is to be carried out without any of the reducible substance present, and to subtract the "blank" so obtained from the apparent diffusion current (65, 66). It is possible that the proportionately larger values of i_d at small concentrations, found by Hamamoto and Maas, were due to the fact that they did not adequately correct for the residual current.

Even when maxima are satisfactorily eliminated and proper correction is made for the residual current, there are cases where the corrected diffusion current is poorly defined and changes with the potential. These unfavorable cases are most frequently encountered in irreversible reductions (reduction of uranyl ion and of several organic compounds), and they require much more study before satisfactory analytical procedures can be developed.

5. Direct test of the Ilković equation

In order to obtain data for a direct test of the Ilković equation, Lingane and Kolthoff (66) measured the diffusion currents of thallous, lead, cadmium, zinc, iodate, and ferricyanide ions over a concentration range from 10^{-4} to 10^{-2} molar in 0.1 *N* potassium chloride containing a trace of methyl red to eliminate the maxima. The (corrected) diffusion currents were in all cases directly proportional to the concentration of the reducible ions. In table 4 are given the experimentally determined diffusion current constants of the foregoing ions (obtained with the same capillary under identical conditions), compared with the corresponding values computed from the Ilković equation.

In order to calculate the diffusion current constants of the various ions from the Ilković equation, it was necessary to know their diffusion coefficients in 0.1 *N* potassium chloride. Since such data are not available in the literature, we (66) computed the various diffusion coefficients by means of the relation,

$$D_i = \frac{RT\lambda_i^0}{zF^2} \quad (12)$$

In this equation, which was first derived by Nernst (84), D_i is the characteristic diffusion coefficient (per mole) of an ion when it is free to diffuse independently of other ions, λ_i^0 is the equivalent ionic conductance of the ion at infinite dilution, z is its valence without regard to sign, and R , T , and F are the gas constant per mole, the absolute temperature, and the faraday, respectively. The fundamental principles underlying this equation have been discussed in detail by Lingane (65).

It is a well-known principle that the ions of a diffusing salt can diffuse more or less independently if a large excess of some other salt is present at equal concentrations on both sides of the diffusion layer (1, 59). Since polarographic diffusion currents are measured with a relatively small concentration of the reducible ions in a solution containing a large excess

TABLE 4
Comparison of the experimentally determined and the calculated diffusion current constants of several reducible ions
In 0.1 N potassium chloride at 25°C.

| ION | $m^{2/3} t^{1/6}$ | K | | DIFFERENCE per cent |
|-----------------------------------------|-------------------------|-----------------------------------------|------------|------------------------|
| | | Observed | Calculated | |
| | $mg.^{2/3} sec.^{-1/2}$ | microamperes per millimole per liter | | |
| Tl ⁺ | 2.28 | 6.13 | 6.17 | -0.7 |
| Pb ⁺⁺ | 2.28 | 8.78 | 8.67 | +1.3 |
| Cd ⁺⁺ | 2.28 | 8.00 | 7.40 | +8.1 |
| Zn ⁺⁺ | 2.24 | 7.65 | 7.27 | +5.0 |
| IO ₃ ⁻ | 2.18 | 26.3 | 26.2 | +0.4 |
| Fe(CN) ₆ ⁴⁻ | 2.28 | 3.78 | 4.12 | -8.3 |

* The data for $m^{2/3} t^{1/6}$ are given for the potentials at which the diffusion currents were measured.

of an indifferent salt, the reducible ions are more or less free to diffuse independently of the other ions in the solution, and equation 12 can therefore be used to compute their diffusion coefficients.

It should be noted that equation 12 is based on the assumption that the diffusing ions behave as ideal solutes, and therefore it is strictly valid only at infinite dilution. In any actual solution the diffusing ions will be subject to interionic forces, which cause them to deviate more or less from ideal behavior. It might be thought that the effect of interionic forces could be corrected for in equation 12 by simply employing the value of the equivalent ionic conductance of the ion in question at the particular ionic strength of the solution (71, 72). However, this results in an over-correction, because, as McBain and Liu (79) have pointed out, the effect of interionic forces is much smaller in the process of diffusion than it is

in the more complicated process of electrolytic conduction. It is an experimental fact that the diffusion coefficient of a salt decreases relatively much less than its equivalent conductance with increasing ionic strength (79). Since there is no better alternative at present, we assumed, in interpreting our experimental results, that the various reducible ions behaved as ideal solutes, and that their diffusion coefficients in 0.1*N* potassium chloride were the same as at infinite dilution.

For convenient reference the diffusion coefficients of various ions at infinite dilution are given in table 5. When the numerical values of *R*

TABLE 5
Calculated diffusion coefficients at infinite dilution of various ions at 25°C.

| ION | λ^0 ohm ⁻¹ cm. ² equiv. ⁻¹ | D cm. ² sec. ⁻¹ $\times 10^6$ | ION | λ^0 ohm ⁻¹ cm. ² equiv. ⁻¹ | D cm. ² sec. ⁻¹ $\times 10^6$ |
|------------------------|---------------------------------------------------------------------------|-------------------------------------------------------------|-----------------------------------------|---------------------------------------------------------------------------|-------------------------------------------------------------|
| H ⁺ | 350 | 9.34 | Ni ⁺⁺ | 52 | 0.69 |
| Na ⁺ | 50.5 | 1.35 | OH ⁻ | 196 | 5.23 |
| K ⁺ | 74 | 1.98 | Cl ⁻ | 76 | 2.03 |
| Li ⁺ | 39 | 1.04 | NO ₃ ⁻ | 72 | 1.92 |
| Tl ⁺ | 75 | 2.00 | CH ₃ COO ⁻ | 41 | 1.09 |
| Cs ⁺ | 79 | 2.11 | IO ₃ ⁻ | 41 | 1.09 |
| Pb ⁺⁺ | 73 | 0.98 | BrO ₃ ⁻ | 54 | 1.44 |
| Cd ⁺⁺ | 54 | 0.72 | SO ₄ ²⁻ | 81 | 1.08 |
| Zn ⁺⁺ | 54 | 0.72 | Fe(CN) ₆ ⁴⁻ | 100 | 0.89 |
| Cu ⁺⁺ | 54 | 0.72 | | | |

(8.317 volt-coulombs), *F* (96,500 coulombs), and *T* (298°K.) are substituted into equation 12, we obtain,

$$D_i = 2.67 \times 10^{-7} \lambda_i^0 / z \text{ cm.}^2 \text{ sec.}^{-1} \quad (13)$$

The diffusion coefficients in table 5 were calculated by means of this equation, using data for the equivalent ionic conductances at infinite dilution taken from the compilation given by MacDougall (70).

The constant 0.627 in equation 7 refers to the current expressed in amperes, the concentration in moles per cubic centimeter, and the product $m^{2/3}t^{1/6}$ in the units gm.^{2/3}sec.^{-1/2}. When the current is expressed in microamperes, the concentration in millimoles per liter, and the product $m^{2/3}t^{1/6}$ in the units mg.^{2/3}sec.^{-1/2}, and the numerical value of *F* is inserted into equation 7, we obtain

$$i_d = 605nD^{1/2}Cm^{2/3}t^{1/6} \quad (14)$$

The diffusion current constant, corresponding to *C* = 1, should therefore be given by

$$K = 605nD^{1/2}m^{2/3}t^{1/6} \quad (15)$$

This equation, with values of D from table 5 and values of the product $m^{2/3}t^{1/6}$ from table 2, was used to compute the calculated values of K in table 4. For example, from table 5 the diffusion coefficient of lead ions is 0.98×10^{-6} cm.² sec.⁻¹, from table 2 the product $m^{2/3}t^{1/6}$ at $E_{d..} = -0.7$ volt (at which potential the diffusion current of lead was measured) was $2.28 \text{ mg.}^{2/3} \text{ sec.}^{-1/2}$, and $n = 2$ in the discharge of lead ions. Therefore, the calculated diffusion current constant of lead ions is

$$K = 605 \times 2 \times (0.98 \times 10^{-6})^{1/2} \times 2.28 = 8.67 \text{ microamperes per millimole per liter}$$

which may be compared with the experimentally determined value, 8.78, given in table 4.

It may be mentioned that the only other direct tests of the Ilkovič equation to be found in the polarographic literature are in his original paper (53). Ilkovič tested his equation in only two cases: in the discharge of 0.001 M cadmium ions from 0.1 N sulfuric acid and in the discharge of 0.001 N hydrogen ions from 0.1 N potassium chloride. He found that the observed diffusion current of cadmium ions was 9 per cent smaller than the value he calculated. However, in calculating the diffusion coefficient of the cadmium ions he overlooked the valence in equation 12 and obtained the incorrect value 1.20×10^{-6} cm.² sec.⁻¹. Hence his calculated value of the diffusion current of cadmium ions was too large.

In view of the complexity of the diffusion processes at the dropping electrode, the agreement between the observed and calculated values of K in table 4 is really remarkably good. It is evident that the Ilkovič equation, and the postulates on which it is based, must be essentially correct. It furnishes a sound theoretical basis for quantitative polarography, and Ilkovič deserves much credit for its derivation.

6. The "condenscr current" and its influence in the measurement of very small diffusion currents

It has already been mentioned that a small current, known as the residual current, is observed before the discharge of a reducible substance begins. Even in the electrolysis of very pure air-free solutions of difficultly reducible salts, such as potassium chloride, an appreciable residual current is found long before the discharge of the ions of the indifferent salt, e.g., potassium ions, finally begins (table 2). The residual current increases in approximately direct proportion to the applied E.M.F., and, as has been pointed out in the preceding section, it must be subtracted from the total observed diffusion current of a given substance in order to obtain the true diffusion current. This correction often amounts to several tenths of a microampere, and it becomes greater the more negative the potential at which the substance in question is reduced (table 2).

It is obvious, therefore, that the accuracy with which small diffusion currents can be measured is greatly dependent on the accuracy of the correction for the residual current. It is also evident that the minimum detectable concentration of a reducible substance will depend on the accuracy with which the residual current can be measured. For example, if the uncertainty in the measurement of the residual current is ± 0.05 microampere, and if the diffusion current constant of a given substance is 10 microamperes per millimole per liter, then the minimum concentration of this substance that could possibly be detected would be of the order of $10^{-5} M$, and would probably have to be somewhat greater in order to detect the substance with certainty. An exact knowledge of the residual current is therefore of great importance when small concentrations of reducible substances are being determined.

In a pure solution of an indifferent salt, that is, in the complete absence of reducible or depolarizing substances, the residual current is due to a so-called "charging current," or "condenser current." It is well known that when a metal, or in fact almost any solid, is placed in an electrolyte solution, an electrical double layer of positively and negatively charged ions establishes itself at the solid-solution interface. This double layer is analogous to an ordinary condenser. The orientation of the charges during its formation causes a small current to flow across the interface, similar to the instantaneous current that flows across an ordinary condenser when it is being charged or discharged. Since the surface of the dropping electrode increases continuously as the mercury drops form, the double layer is being formed continuously on each new element of the freshly forming surface; hence a small continuous charging current results. This is the "condenser current" (54).

As is well known, the change in the strength and the sign of the double layer at a mercury solution interface is responsible for the parabolic shape of the electrocapillary curve of mercury (15). At potentials more positive than correspond to the maximum in the electrocapillary curve the mercury side of the double layer is apparently positively charged, while at potentials more negative than the electrocapillary maximum it is negatively charged. At the maximum itself there is presumably no double layer at the mercury-solution interface. It follows, therefore, that the direction in which the condenser current flows, i.e., from the mercury to the solution, or *vice versa*, should depend on the sign of the charge on the mercury surface and that it should change direction at the electrocapillary maximum. Since there is no double layer at the electrocapillary maximum, the condenser current should be zero at this point. These predictions are verified by the residual current curves shown in figure 11. These curves were obtained by Maas (68) by electrolyzing a 0.1 N solution of

pure sodium chloride from which the air had been displaced by purified hydrogen. The curves were recorded with a polarograph, using the full sensitivity of the galvanometer.

Curve 1 in this polarogram is practically a "pure" condenser current, and it is seen that it is approximately directly proportional to the potential of the dropping electrode up to about -1.6 volts, where the discharge of sodium ions begins. The electrocapillary maximum in 0.1 N chloride solution is at about -0.56 volt, so the condenser current should have been zero at this potential. Actually it is seen that the current was zero at a somewhat more positive potential, at about -0.35 volt. Up to -0.35 volt the current was negative, that is, positive electricity flowed from the

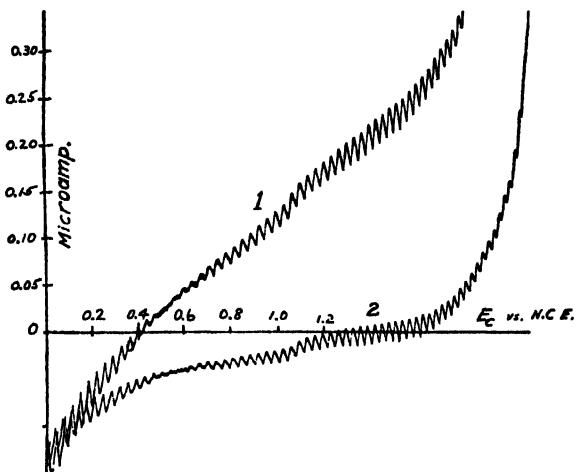


FIG. 11. Condenser current curves in pure 0.1 N sodium chloride according to Maas (obtained with polarograph).

dropping electrode to the solution, whereas above -0.35 volt positive electricity flowed from the solution to the electrode. It will be noted that the current oscillations during the formation of the individual mercury drops decreased with increasing negative potential and finally practically disappeared between -0.45 and -0.55 volt. As the potential was increased above -0.55 volt, the oscillations reappeared and their magnitude increased with increasing negative potential. The current oscillations are an approximate measure of the magnitude of the charging current during the formation of the individual mercury drops. In agreement with the theory of the electrocapillary curve, the minimum in the oscillations occurred at the potential corresponding to the maximum in the electrocapillary curve, where the double layer disappears (68, 124).

It should be mentioned that Ilkovič (55) measured the potential-time curves of individual mercury drops falling into a very pure, air-free solution of 0.1 *N* potassium chloride. At potentials more positive than the electrocapillary maximum he found that the potential of an individual drop becomes more positive during its growth. At the potential of the electrocapillary maximum the potential remained practically constant during the life of a drop, while at more negative potentials it became more negative during the life of each drop. These results confirm the idea that the double layer changes sign at the electrocapillary maximum and that no double layer exists at the maximum itself.

It will be noted from curve 1 in figure 11 that the residual current was slightly greater than zero (0.03 microampere) at the maximum in the electrocapillary curve (-0.56 volt). This was probably due to the reduction of a trace of oxygen remaining in the solution, because it is practically impossible to remove the very last traces of oxygen from a solution by simply bubbling hydrogen or nitrogen through it (68). In most cases a residual current will include a small "faradayic" current due to the reduction of traces of reducible impurities in the solution. That is,

$$i_r = i_e + i_f \quad (16)$$

where i_r is the total residual current, i_e is the condenser current, and i_f is the faradayic current.

If the double layer behaves as a condenser, the *average* charging current during the life of a drop should be given by

$$i_e = kq\psi \quad (17)$$

where k is the capacity of the double layer, q is the *average* rate at which fresh mercury surface is being exposed to the solution (cm.^2 per second) during the life of a drop, and ψ is the average potential difference between the mercury and the solution (55, 68). Since i_e is zero at the potential corresponding to the electrocapillary maximum, and since k remains fairly constant for small changes in the potential, equation 17 can be written as

$$i_e = kq(\psi_{\max.} - \psi) = kq(E_{\max.} - E) \quad (18)$$

where $\psi_{\max.}$ is the absolute potential difference between the mercury and the solution at the electrocapillary maximum, and i_e is the condenser current at some other potential ψ . Since the potential of the quiet electrode remains constant during the electrolysis, and since the potential of the solution is a constant, $(\psi_{\max.} - \psi)$ can be replaced by $(E_{\max.} - E)$, where E is the potential of the dropping electrode with respect to an external reference electrode. Since q and k can be measured, this equation allows us to calculate i_e at various potentials of the dropping electrode.

Ilković (55) calculated the capacity of the double layer in specially purified 0.1 *N* potassium chloride from his measurements of the potential-time curves of individual mercury drops. On the negative branch of the electrocapillary curve (-1.3 volts) he found a value for *k* of 22.3 microfarads per square centimeter, while on the positive side of the electrocapillary maximum (-0.18 volt) he found a value for *k* of 42.2 microfarads per square centimeter. These results are in agreement with the measurements of Philpot (92), who found values for *k* of 21.8 and 48.8 microfarads per square centimeter on the negative and the positive side of the electrocapillary maximum, respectively. In this connection important investigations by Proskurnin and Frumkin (96) of the capacity of the double layer in solutions of sodium sulfate, potassium chloride, sodium hydroxide, and sulfuric acid should be mentioned.

It is interesting to compare the value of *i_e* computed by means of equation 18 with the data for the residual current in 0.1 *N* potassium chloride given in table 2. It can be derived that the area exposed per second by the growing mercury drops is given by

$$q = (4\pi)^{1/3}(3/d)^{2/3}m^{2/3}t^{-1/3} = 0.0085m^{2/3}t^{-1/3} \text{ cm.}^2 \text{ per second} \quad (19)$$

where *d* is the density of mercury, *t* is the drop time, and *m* is expressed in milligrams per second. From the data in table 2 we find that *m* was 2.67 mg. per second, and *t* was 2.18 sec. at -1.5 volts, and hence *q* was 0.0126 cm.² per second. Since *E*_{max.} is equal to -0.56 volt with respect to the normal calomel electrode, and *k* is equal to 22 microfarads per square centimeter, we calculate for *i_e* at -1.5 volts,

$$i_e = 22 \times 0.0126 \times (-0.56 + 1.5) = 0.28 \text{ microampere} \quad (20)$$

This agrees fairly well with the experimentally observed value of 0.38 microampere in table 2. It should be mentioned that the change in ψ during the life of a drop has been neglected in deriving equation 18, because it cannot be calculated theoretically; hence the calculated value of *i_e* is probably somewhat too small. It is evident, however, that about three-fourths of the residual current in table 2 is a condenser current.

The capacity of the double layer can also be computed approximately from the slope of the condenser current curve. From equation 18 we see that the capacity is given by

$$k = (di_e/dE)/q \quad (21)$$

that is, *k* is equal to the slope of the condenser current curve divided by the average area exposed per second. Even if the residual current includes a small faradayic current, it may be assumed that the latter will be practically constant over a small range of potentials, and the slope of the residual current curve will be practically equal to *di_e/dE*. For example,

the slope of curve 1 in figure 11 is approximately 0.20 microampere per volt, and q in this experiment of Maas was 0.0086 cm.^2 per second (68). The value of k calculated from these data is $0.20/0.0086 = 23$ microfarads per square centimeter, which is in reasonably good agreement with the value 22 found by Ilković and Philpot by different methods. Since the condenser current curve in figure 11 is approximately a straight line at potentials more negative than the electrocapillary maximum, it is evident that the capacity of the double layer is approximately constant between about -0.4 volt and -1.4 volts. It will be noted that the slope of the curve (and hence the capacity of the double layer) is larger on the positive side of the electrocapillary maximum, in agreement with the results of Ilković, Philpot, and Proskurnin and Frumkin.

The capacity of the double layer (and hence the magnitude of the condenser current) depends on the nature of the ions in the solution. Capillary-active ions which are strongly adsorbed at the mercury-solution interface increase the capacity of the double layer on one or the other branch of the electrocapillary curve. For example, the capacity of the double layer is about twice as great in a solution containing iodide ions as it is in a chloride solution (on the positive side of the electrocapillary curve) (15). This corresponds to a decrease in the thickness, or a compression, of the double layer acting as a condenser. The effect of capillary-active non-electrolytes (octyl alcohol) has been investigated by Proskurnin and Frumkin (96).

In order to eliminate as far as possible the uncertainty in the measurement of very small diffusion currents, due to the residual current, Ilković and Semerano (57) recommend that the condenser current be "compensated" by sending a current of equal magnitude through the current-measuring galvanometer in a direction opposite to that of the condenser current. Recent models of the polarograph are equipped with a device for automatically compensating, or balancing out, the residual current, at all values of the applied E.M.F. (50, 68). Curve 2 in figure 11 was obtained by Maas (68) by compensating the residual current in this way.

7. The influence of temperature on the diffusion current

Ilković (56) has recently shown that the effect of temperature on the diffusion current can be readily interpreted as due chiefly to the change in the diffusion coefficient of the reducible substance with temperature. We have seen that the diffusion coefficient of an ion is directly proportional to both the absolute temperature and the equivalent conductance of the ion at infinite dilution. The equivalent conductances of most ions increase by about 2 to 2.5 per cent per degree. At $T = 298^\circ\text{K}$. an increase of one

degree will increase the value of T by about 0.3 per cent. Furthermore, m increases by about 0.5 per cent per degree, owing to the decrease in the viscosity of mercury with increasing temperature (63, 56). Hence, from the cumulative effect of these three factors, it is to be expected that the diffusion current will increase by about 1.6 per cent per degree increase in temperature (56). This prediction is in reasonably good agreement with the experimental results of Nejedlý (83), Kemula (61), and Majer (73), who found that the temperature coefficient of the diffusion current was about 2 per cent per degree for most metal ions.

From the analytical viewpoint it is evident that the temperature should be controlled to at least $\pm 0.5^{\circ}\text{C}$., or better, in order to keep errors due to the temperature effect within ± 1 per cent.

8. Influence of the solvent on the diffusion current constant

The influence of a change in the nature of the solvent on the diffusion current constant of a given reducible substance can be predicted from the Ilkovič equation. According to equation 15 the diffusion current constant, K , depends on the quantities D , m , and t . We have shown that m for a given capillary at a given pressure on the dropping mercury and a constant temperature is practically independent of the medium in which the drops form. The drop time, being proportional to the interfacial tension, will vary with the nature of the medium, but, since K depends on only the sixth root of t , considerable variations in the latter will have only a relatively small effect on K . Therefore, with the same capillary at a given temperature and pressure, it is evident that any change in K in going from water to some non-aqueous solvent will be due chiefly to a change in the diffusion coefficient of the reducible substance. In general, the diffusion coefficients of most reducible substances are smaller in non-aqueous solvents than in water. It is to be expected, therefore, that the diffusion current constants of most substances will be smaller in non-aqueous solvents than in water.

Perrachio and Meloche (89) have recently studied the diffusion currents of the various alkali ions in alcohol-water, dioxane-water, ethylene glycol-water, and glycerol-water mixtures. They found that the diffusion currents were smaller in these solvents than in pure water. The diffusion currents in the glycerol-water mixtures were the smallest of all; this is to be expected because of the relatively great viscosity and correspondingly small diffusion coefficients in these mixtures. Quantitative interpretation of their results is not possible because the requisite diffusion coefficient data are not available. Qualitatively, however, their data are in accord with the foregoing predictions.

VII. THE MIGRATION CURRENT AND ITS INFLUENCE ON THE
LIMITING CURRENT

1. Relation between limiting currents in the absence of added salt and diffusion currents in the presence of an excess of indifferent salt

In the preceding section we considered the characteristics of the diffusion current, which is the special name applied to the limiting current when an excess of some indifferent salt is present in the solution. If there is no other salt present, the current will be carried entirely by the ions of the reducible salt, and therefore the reducible ions will be subject to electrical

TABLE 6

Influence of various concentrations of potassium chloride, potassium nitrate, and hydrochloric acid on the limiting current of lead chloride

50 cc. of 0.00095 M lead chloride plus 0.2 cc. of 0.1 per cent sodium methyl red, with various concentrations of added electrolytes. Temperature = 25°C.,
 $P = 21.8 \text{ cm.}, m^{2/3}t^{1/6} = 2.28 \text{ mg.}^{2/3} \text{ sec.}^{-1/2}$

| ADDED SALT <i>equivalents per liter</i> | ¹ | | |
|--------------------------------------------|----------------------------|-----------------------------------------|----------------------------|
| | KCl <i>microamperes</i> | KNO ₃ <i>microamperes</i> | HCl <i>microamperes</i> |
| 0 | 17.6 | 17.6 | 17.6 |
| 0.0001 | 16.3 | 16.2 | 15.7 |
| 0.0002 | 14.9 | 15.0 | 14.6 |
| 0.0005 | 13.3 | 13.4 | 12.7 |
| 0.001 | 11.8 | 12.0 | 11.2 |
| 0.005 | 9.8 | 9.8 | 9.5 |
| 0.1 | 8.35* | 8.45* | † |
| 1.0 | 8.00* | 8.45* | † |

* Corrected for the residual current.

† Maxima present, and diffusion currents not well defined.

migration as well as to diffusion. Under these conditions the limiting current will evidently depend on the amount of reducible ions supplied by electrical migration as well as by diffusion. In this section we shall consider the effect of this electrical migration on the limiting current, and also the relation between limiting currents without added salt and diffusion currents in an excess of indifferent salt.

In the reduction of cations it has been found that the limiting current without added salt is roughly twice as great as the diffusion current obtained with an excess of indifferent salt in the solution. As indifferent salt is added to the pure solution of a salt of a reducible cation, the limiting current decreases very rapidly with the first small additions and then more slowly as the concentration of indifferent salt is increased, until it finally

becomes practically constant and independent of further additions of the indifferent salt when the concentration of the latter has been made about fifty times larger than that of the reducible salt. Several examples of this behavior have been given by Šlendyk (110) and Lingane (65). A typical example is given in table 6 (65).

It has been found that anion limiting currents without any added salt are *increased* by the addition of indifferent salt. For example, we (65) found a limiting current of 2.40 microamperes for a pure solution of $1.5 \times 10^{-4} M$ potassium iodate (reduction of iodate), which increased to 3.70 microamperes when the solution was made 0.1 N in potassium chloride (see table 7).

In order to account for these observed effects, Heyrovský (36) and Ilkovič (53) proposed the idea that the limiting current, without added salt or with only a relatively small amount of added salt, was the sum of a diffusion current, i_d , and a migration current, i_m . In the reduction of cations the limiting current should be increased by the electrical migration, because cations will be impelled towards the negatively charged electrode. Hence Heyrovský and Ilkovič write

$$i_l = i_d + i_m \quad (\text{cation reduction}) \quad (22)$$

where i_l is the total limiting current. In the electroreduction of anions the limiting current should be decreased by the electrical migration, because anions will be repelled from the negatively charged electrode, so that in this case we have

$$i_l = i_d - i_m \quad (\text{anion reduction}) \quad (23)$$

This formal division of the limiting current into two parts is based on the concept, already explained, that in the absence of added salt the reducible ions are subject to both a diffusive force and an electrical force. These forces are assumed to act independently of each other; each is assumed to contribute its share to the total flux of reducible ions and hence to the total limiting current. This concept of the independent action of a diffusive force and an electrical force is not new; it was used by Nernst (84), Haskell (23), Planck (94), and various other investigators, and it is incorporated in the modern theory of electrolytic diffusion and the diffusion potential (87, 30, 22).

A second postulate made by Heyrovský is that the migration current is given by the product of the *total* limiting current and the transference number T_i of the reducible ions, that is,

$$i_m = T_i i_l \quad (24)$$

When no other salt is present, the transference number in this equation will be designated thus, T_+^0 , to indicate that it is the transference number of the reducible ion in the pure solution of the reducible salt. When extraneous salt is added to the solution, the transference number will be decreased from T_+^0 to some value T_+ , which will depend on the relative concentrations and the mobilities of the various ions. As extraneous salt is added to the solution, T_+ decreases very rapidly.

Heyrovský and Ilkovič assume that the diffusion current, i_d , is the same in the pure solution of the reducible salt as in the presence of an excess of

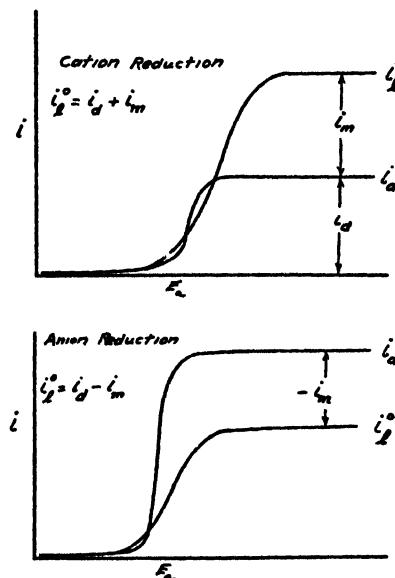


FIG. 12. Relation between limiting currents without added salt and diffusion currents with an excess of added salt (schematic).

added salt. On the basis of this assumption we find, by combining equations 22 and 23 with equation 24,

$$\frac{i_2^0}{i_d} = 1/(1 - T_+^0) \quad (\text{cation reduction}) \quad (25)$$

$$\frac{i_2^0}{i_d} = 1/(1 + T_-^0) \quad (\text{anion reduction}) \quad (26)$$

In these equations i_2^0 is the initial limiting current without any foreign salt in the solution. Since the transference numbers of most reducible ions are in the neighborhood of 0.5, it is evident that the limiting current in most cation reductions should be decreased to about one-half of its initial value by adding an excess of foreign salt to the solution, whereas the initial limiting current in most anion reductions should be increased

to about one and one-half times its initial value. These relations are shown schematically in figure 12 for the case in which $T^0 = 0.5$

In deriving the foregoing relations Heyrovský assumed that the diffusion component of the limiting current was constant and independent of the presence of foreign salt. However, MacGillavry (71) and Lingane (65) have shown that this is only an approximation to the facts, because the diffusion component of the limiting current changes appreciably when foreign salt is added to the solution. According to the Ilkovič equation (see preceding section), with all other factors constant,

$$i_d = kD^{1/2} \quad (27)$$

When an excess of foreign salt is present in the solution, the diffusion coefficient in this equation will be the characteristic diffusion coefficient of the reducible ion (equation 12). However, when no foreign salt is present, the reducible ions are not free to diffuse according to their characteristic diffusion coefficient, but must diffuse at a rate proportional to the diffusion coefficient of the reducible salt itself. It is readily derived (23, 30, 65, 87, 94, 116), that the diffusion coefficient of an ideal salt is given by

$$D_{\text{salt}} = \frac{RT}{F^2} \left(\frac{\lambda_+^0 \lambda_-^0}{\lambda_+^0 + \lambda_-^0} \right) \left(\frac{1}{z_+} + \frac{1}{z_-} \right) = \frac{D_+ D_- (z_+ + z_-)}{z_+ D_+ + z_- D_-} \quad (28)$$

where λ_+^0 and λ_-^0 are the equivalent conductances of the positive and negative ions of the salt at infinite dilution, z_+ and z_- are the valences, and D_+ and D_- are the characteristic diffusion coefficients of the ions. This equation, like equation 12, has been derived on the assumption that the ions behave as ideal solutes, and it therefore will be strictly correct only at infinite dilution. However, no appreciable error should result in applying it to pure solutions of reducible salts, when the ionic strength is small. In general, D_{ion} and D_{salt} will not be equal (except in the special case when $\lambda_+^0 = \lambda_-^0$ and $z_+ = z_-$), and hence it is evident from equation 27 that the diffusion component of the limiting current in the pure solution of a reducible salt will in general be different from the diffusion current in an excess of added salt.

When an excess of foreign salt is present, the diffusion coefficient in equation 27 will be D_{ion} , given by equation 12, whereas in a pure solution of a reducible salt the diffusion coefficient will be D_{salt} , given by equation 28. It is evident, therefore, that

$$i_d^0 = kD_{\text{salt}}^{1/2} \quad (29)$$

$$i_d = kD_{\text{ion}}^{1/2} \quad (30)$$

where i_d^0 is the diffusion component of the limiting current without added salt, and i_d is the diffusion current with an excess of foreign salt present. From equation 7 we see that k in the foregoing equations is given by

$$k = 0.627nFCm^{2/3}t^{1/6} \quad (31)$$

The first four terms in this expression are obviously the same with or without indifferent salt present in the solution, and we have shown in the preceding section that the product $m^{2/3}t^{1/6}$ is practically independent of the composition of the solution. It is evident, therefore, that k will be the same with and without foreign salt present, and that the change in i_d as foreign salt is added to the solution is almost entirely due to the change in the effective diffusion coefficient.

By combining equations 29 and 30 with equations 22, 23, and 24, we obtain

$$\frac{i_l^0}{i_d} = (D_{\text{salt}}/D_{\text{ion}})^{1/2}[1/(1 - T_+^0)] \quad (\text{cation reduction}) \quad (32)$$

$$\frac{i_l^0}{i_d} = (D_{\text{salt}}/D_{\text{ion}})^{1/2}[1/(1 + T_-^0)] \quad (\text{anion reduction}) \quad (33)$$

These expressions should be more accurate than the simpler equations 25 and 26 derived by Heyrovský.

MacGillavry (71) has also derived an equation for the ratio i_l^0/i_d in cation reductions. In the notation of this paper his equation is

$$\frac{i_l^0}{i_d} = \left(1 + \frac{z_+}{z_-}\right)^{1/2} \left(1 + \frac{z_-\delta_+^0}{z_+\delta_-^0}\right)^{1/2} \left(\frac{\delta_+^0}{\delta_+}\right)^{1/2} \left(\frac{t^0}{t}\right)^{1/6} \quad (34)$$

where δ_+^0 and δ_-^0 are the absolute mobilities at infinite dilution of the cation and anion of the reducible salt, δ_+ is the absolute mobility of the reducible cation in the presence of an excess of added salt, and t^0 and t are the drop times, at constant pressure, without and with added salt. Since

$$\delta_i^0 = \lambda_i^0 / zF^2$$

and

$$D_{\text{ion}} = RT\lambda_i^0 / zF^2$$

it can be shown that the *first two* terms in MacGillavry's equation are identical with equation 32. The third term is a correction for the decrease in the absolute mobility of the reducible cation (salt effect) when foreign salt is added to the solution, and the last term is a correction for the change in drop time when foreign salt is added. We have already shown that the change in t as indifferent salt is added is small, unless the added salt is strongly capillary-active, and therefore the ratio t^0/t will generally be close to unity. Since only the sixth root of this ratio is involved, the

last term will be so close to unity in most cases that it can be safely neglected.

Formally then, the essential difference between MacGillavry's equation and equation 32 is the correction, $(\delta_+^0/\delta_+)^{1/2}$, for the decrease in the absolute mobility of the reducible cation as indifferent salt is added to the solution. However, this difference is more apparent than real, because if the true (measured) values of D_{ion} and D_{salt} are used in equations 32 and 33, these equations become theoretically just as exact as equation 34. Unfortunately, we have no exact knowledge of the decrease in the absolute mobility of an individual ion with increasing ionic strength. Qualitatively, there is good reason for believing that the mobility of an ion decreases much less rapidly than its equivalent conductance with increasing ionic strength. The equivalent conductance of an ion is a true measure of its mobility only at infinite dilution but not at any finite concentration (79, 87). Furthermore, there are practically no reliable diffusion coefficient data in the literature for the salts and ions with which we are concerned, at various ionic strengths. For these reasons we shall assume, in the following interpretation of experimental data, that the various salts and ions behave as ideal solutes.

In order to test the foregoing relations, we (65) determined the initial limiting currents, with no added salt, and the diffusion currents, in 0.1 *N* or 0.9 *N* potassium chloride, of various thallous salts, lead chloride, cadmium sulfate, and potassium iodate. The various thallous salts were chosen so that the transference number of the thallous ion would differ as much as possible in the various cases. A trace of methyl red (one drop of a 0.1 per cent solution of sodium methyl red per 50 cc.) was added to the solutions to eliminate the maxima otherwise present. Special experiments showed that this very small amount of methyl red had no appreciable influence on either the limiting or the diffusion currents. The results obtained are given in table 7.

The transference numbers of the reducible ions in the pure solutions of the reducible salts, given in the fourth column, were calculated in the usual way from the equivalent conductances of the ions of the reducible salt. The fifth column of the table contains the observed ratios of the initial limiting currents to the final diffusion currents, i_l^0/i_d . The values of this ratio computed according to the simple Heyrovský relation (equation 25 or 26) are tabulated in the sixth column, while the last column contains the values of i_l^0/i_d computed according to equation 32 or equation 33.

It will be noted that the values of i_l^0/i_d calculated according to equation 32 or equation 33 are, in most cases, in better agreement with the observed values of this ratio than the values calculated according to the simpler equations of Heyrovský (equation 25 or equation 26). Cadmium sulfate

and potassium iodate are the only exceptions; in these cases, equation 25 or equation 26 gave better results.

It will be noted that the observed values of the ratio are, in most cases, less than the calculated values. This is perhaps due to unavoidable traces of salt impurities in the pure solutions of the reducible salts, which would decrease i_i^0 appreciably, e.g., traces of alkaline impurities from the glass of the cell, or traces of dissolved mercury from the quiet electrode. In view of the difficulty of completely eliminating such salt impurities, the agreement between the observed and calculated ratios is as good as can be reasonably expected.

TABLE 7

Comparison of initial limiting currents and diffusion currents of various reducible ions

| REDUCIBLE SALT | i_i^0 microamperes | i_d micro-amperes | T_i^0 (CALCULATED) | RATIO i_i^0/i_d | | |
|--------------------------------------------------------------------|--------------------------------------|------------------------|---------------------------------------|-------------------|---------------------------|---------------------------|
| | | | | Observed | Calculated by equation 25 | Calculated by equation 32 |
| 0.001 M TlCl | 11.6 | 6.10* | 0.500 | 1.90 | 2.00 | 2.00 |
| | 11.6 | 5.95† | 0.500 | 1.95 | 2.00 | 2.00 |
| 0.0013 M TlOH | 13.9 | 8.15* | 0.277 | 1.71 | 1.39 | 1.66 |
| 0.0026 M TlOH | 26.7 | 15.85* | 0.277 | 1.69 | 1.39 | 1.66 |
| 0.0026 M TlCl ₂ H ₂ O ₂ | 34.2 | 15.85* | 0.646 | 2.16 | 2.83 | 2.37 |
| 0.0005 M TlIO ₃ | 6.55 (Tl ⁺) | 3.30* | 0.646 (Tl ⁺) | 1.99 | 2.83 | 2.37 |
| 0.00095 M PbCl ₂ | 17.6 | 8.35* | 0.483 | 2.11 | 1.93 | 2.39 |
| | 17.6 | 8.00† | 0.483 | 2.20 | 1.93 | 2.39 |
| 0.001 M CdSO ₄ | 12.7 | 8.00* | 0.400 | 1.59 | 1.67 | 1.83 |
| 0.00015 M KIO ₃ | 2.40 (IO ₃ ⁻) | 3.70* | 0.356 (IO ₃ ⁻) | 0.65 | 0.74‡ | 0.84§ |

* In 0.1 N potassium chloride corrected for the residual current.

† In 0.9 N potassium chloride corrected for the residual current.

‡ Calculated by equation 26.

§ Calculated by equation 33.

It is interesting to note that a correction for the decrease in the absolute mobilities of the reducible cations when foreign salt is added, as recommended by MacGillavry (equation 34) (71), would increase the calculated ratios of i_i^0/i_d . Hence such a correction would actually increase the differences between the observed and calculated values of this ratio. The foregoing experimental data indicate that such a correction is not justified, and that other unknown factors (specific salt effects, etc.) are operative to a certain degree. The decrease in the diffusion current of lead ions between 0.1 N and 1 N potassium chloride (table 6) is an example of such a specific salt effect (probably PbCl₅⁻ or PbCl₄²⁻ formation).

It should be mentioned that the limiting currents of uncharged (especially organic) substances should be practically entirely diffusion currents, because uncharged substances are not subject to electrical migration. The limiting currents of uncharged substances should therefore be constant and independent of the presence of foreign salts (53). Experimentally, however, this is difficult to demonstrate, because the reduction of most uncharged (organic) substances involves the action of hydrogen ions, and, if no foreign acid is present, the necessary hydrogen ions must be furnished by the dissociation of water molecules. The reduction of uncharged substances is thus a much more complicated problem than the reduction of ions (80). It has been found, for example, that the limiting current of oxygen in potassium chloride solutions is practically constant and independent of the concentration of the salt up to about 0.1 N (50, 124), but that in very concentrated salt solutions the limiting current of oxygen is greatly decreased (50).

2. Increase, or "exaltation," of the migration current by the preceding discharge of an uncharged substance

We have seen that the migration current is equal to the product of the total limiting current and the transference number of the reducible ions (equation 24). If some uncharged substance, which is reduced at a more positive potential than the reducible ion in question, is added to a pure solution of a reducible salt, the resulting total limiting current will be the sum of that due to the reduction of the uncharged substance and that of the reducible ions. That is,

$$i_{\text{total}} = i_u + i_i \quad (35)$$

where i_u is the limiting current of the uncharged substance and i_i is that of the reducible ions. Under these conditions the migration current of the reducible ions will be increased or "exalted" (42), and if we assume that the transference number of the reducible ions is not influenced by the preceding discharge of the uncharged substance, the exalted migration current should be given by

$$i_m = (i_u + i_i) T_i^0 \quad (36)$$

The effect of this exaltation of the migration current will be to increase the limiting current in cation reductions and to decrease it in anion reductions. This effect has been demonstrated by Heyrovský and Bureš (42), who found that the initial limiting currents of potassium and sodium ions in very dilute, pure solutions of the alkali chlorides were increased by the preceding discharge of oxygen. This effect of oxygen is shown by the curves in figure 13, which we (65) have obtained. Curve 1 in this figure

was obtained by electrolyzing a pure $9.4 \times 10^{-4} N$ solution of potassium chloride, from which the air had been completely removed with nitrogen. Curve 2 was obtained after saturating the solution with pure oxygen. It is seen that the limiting current of the potassium ions was increased more than twofold by the preceding discharge of oxygen.

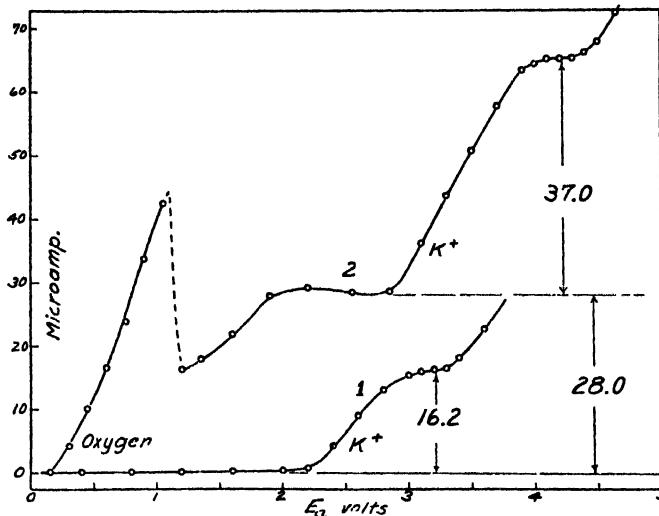


FIG. 13. Exaltation of the limiting current of potassium ions by the preceding discharge of oxygen.

If we represent the initial limiting current, in the absence of foreign salts and without a reducible non-electrolyte present, by i_l^0 and the migration current under the same conditions by i_m^0 , then

$$i_m^0 = i_l^0 T_i^0 \quad (37)$$

Hence the exaltation of the migration current by the preceding discharge of the uncharged substance should be given by

$$i_m - i_m^0 = (i_u + i_l - i_l^0) T_i^0 \quad (38)$$

We also have the relations

$$i_l^0 = i_d \pm i_l^0 T_i^0 \quad (39)$$

$$i_l = i_d \pm (i_u + i_l) T_i^0 \quad (40)$$

where the plus sign refers to cation reductions and the minus sign to anion reductions. If we assume that i_d is unaffected by the preceding discharge of the uncharged substance, we find that the exaltation of the limiting

current in cation reductions should be equal to the exaltation of the migration current and should be given (42) by

$$i_l - i_l^0 = i_m - i_m^0 = i_u T_+^0 / (1 - T_+^e) = i_u \lambda_+^0 / \lambda_-^0 \quad (41)$$

On the other hand, in anion reductions the limiting current should be decreased by the same amount that the migration current is increased. It is interesting to note that in the special case where T_+^0 is exactly 0.5, the exaltation should be simply equal to the limiting current of the uncharged substance, that is,

$$\Delta i_m = i_u$$

when $T_+^0 = 0.5$.

From equation 41 we see that the exaltation of the limiting current should be independent of the concentration of the reducible ions and should depend only on the magnitude of the limiting current of the uncharged substance. Hence it should be possible to increase cation limiting currents many fold by simply increasing the concentration of the uncharged substance and to suppress entirely anion limiting currents. It should be realized, however, that equation 41 will no longer apply with extremely small concentrations of the reducible salt, because under these conditions the electrical potential gradient at the surface of the mercury drops will be so great that the ordinary laws of ionic transfer and conduction will no longer apply.

In curve 2 of figure 13 the limiting current of oxygen, i_u , is 28.0 microamperes. Since the transference number of the potassium ion in potassium chloride is 0.493, we calculate from equation 41 that the exaltation of the potassium ion limiting current by the preceding discharge of oxygen should be $28.0 \times 0.493 / 0.507 = 27.2$ microamperes. The observed exaltation is $37.0 - 16.2 = 20.8$ microamperes.

Similar experiments were carried out with $9.4 \times 10^{-4} N$ sodium chloride. In the absence of oxygen the limiting current of sodium ions was 12.5 microamperes. When the solution was saturated with pure oxygen, the limiting current of sodium ions was 26.5 microamperes, while that of the oxygen was 30.0 microamperes. Hence the observed exaltation was $26.5 - 12.5 = 14.0$ microamperes. Since the transference number of the sodium ion in sodium chloride is 0.400, the calculated exaltation is $30.0 \times 0.400 / 0.600 = 20.0$ microamperes.

Although the discrepancy between the calculated and observed exaltation in the foregoing cases is in the neighborhood of 25 per cent, the agreement is sufficiently good to show that Heyrovsky's theory of the exaltation is at least approximately correct. The discrepancies can be accounted for by the fact that hydroxide ions are a product of the reduc-

tion of oxygen, and, when the latter is discharging simultaneously with the alkali ions, the solution close to the surface of the mercury drops is actually equivalent to a mixture of the alkali chloride and hydroxide. Hence the transference numbers of the alkali ions, in the solution close to the surface of the electrode, are considerably less than in the pure solutions of the alkali chlorides, and the exaltation is correspondingly decreased. For example, in the case of potassium chloride the calculated diffusion current in the absence of oxygen is $(1 - 0.493)16.2 = 8.2$ microamperes, and it was the same when oxygen was discharging. Therefore the observed migration current when oxygen was discharging was $37.0 - 8.2 = 28.8$ microamperes. The effective transference number of the potassium ion during the simultaneous discharge of oxygen, calculated from this value of the migration current, is $28.8/65.0 = 0.443$, instead of the value 0.493 in pure potassium chloride. This corresponds to a ratio of hydroxide to chloride ions of about 1/5 close to the surface of the mercury drops, which is a logical value. Heyrovský and Bureš apparently did not recognize the necessity of correcting for the decrease in the effective transference numbers of the alkali ions due to the accumulation of hydroxide ions at the surface of the mercury drops, probably because they saturated their solutions with air instead of pure oxygen, in which case the correction will be much less.

Heyrovský and Bureš state that they also obtained the exaltation effect with barium and manganous ions and by using other reducible non-electrolytes in place of oxygen, e.g., quinone. The reduction of uncharged substances practically always involves hydrogen ions, which in neutral solutions must be furnished by the dissociation of water molecules at the surface of the mercury drops, with the subsequent accumulation of hydroxide ions. Hence with practically all uncharged substances the formation of hydroxide ions at the electrode surface must be taken into account.

From equation 41 we see that the exaltation of the sodium and potassium ion limiting currents should be directly proportional to the equivalent conductances of the two alkali ions, when the anion in both cases is the same. The exaltation of the potassium ion limiting current will therefore be considerably greater than that of the sodium ion, at a given value of i_u . In a mixture of potassium and sodium chlorides the observed exaltation will be intermediate between the characteristic exaltations of the two alkali ions. Heyrovský and Bureš recommend this behavior for determining the ratio of the two alkali chlorides in a mixture.

3. Exaltation of the migration current by the preceding discharge of a reducible ion

The apparent limiting current of a reducible cation can be increased or exalted by the preceding discharge of a reducible ion, as well as by the

preceding discharge of a reducible non-electrolyte. This effect is shown in figure 14. Curve 1 in this figure was obtained by electrolyzing a $9.4 \times 10^{-4} N$ potassium chloride solution that had been freed from air with pure nitrogen. Curve 2 was obtained after addition of sufficient thallous chloride solution to make the concentration of the latter 8.3×10^{-4} (the concentration of the potassium chloride being reduced to $8.6 \times 10^{-4} N$).

Without thallous chloride present the potassium ion limiting current was 15.8 microamperes, and it was increased to 20.1 microamperes by the preceding discharge of the thallous ions. Without thallous chloride present, the calculated diffusion current of potassium ions was

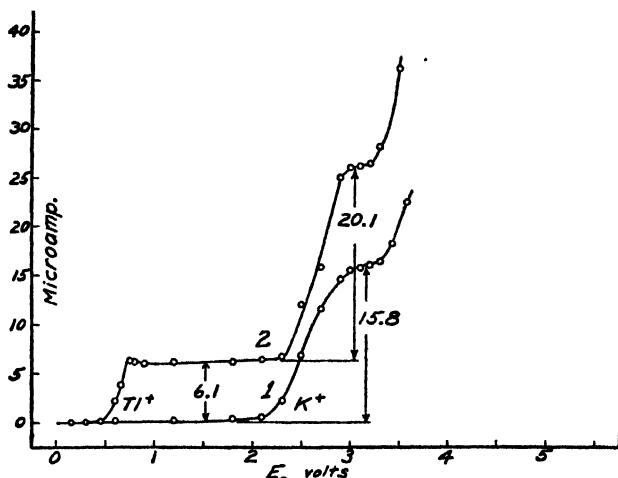


FIG. 14. Exaltation of the limiting current of potassium ions by the preceding discharge of thallous ions.

$15.8(1 - 0.493) = 8.0$ microamperes. Since thallous, potassium, and chloride ions all have practically the same equivalent conductances, the effective diffusion coefficient of the potassium ions will be the same with or without thallous chloride present, but the diffusion current of potassium ions was decreased to $8.0 \times 8.6/9.4 = 7.3$ microamperes by the increase in volume when the thallous chloride solution was added. Hence the migration current when thallous chloride was present was $20.1 - 7.3 = 12.8$ microamperes.

It should be realized that when the potential corresponding to the limiting current of potassium ions was reached, in the presence of thallous chloride, the thallous ion migration current was also increased, as well as that of the potassium ions, by the increased current. Hence only a part of the apparent increase in the potassium ion limiting current was actually

due to the exaltation of the potassium ion limiting current, and the remainder was due to the exaltation of the migration current of the thallous ions during their simultaneous discharge with potassium ions! In the solution containing almost equal concentrations of thallous chloride and potassium chloride, since the equivalent conductances of all the ions are very nearly the same, the transference numbers of the thallous and potassium ions were practically the same and were equal to 0.25, while that of the chloride ion was 0.5. Since the total limiting current was 26.2 microampères, the migration currents of the thallous and the potassium ions should each have been equal to $0.25 \times 26.2 = 6.55$ microampères, and the total migration current should have been equal to 13.1 microampères. This calculated migration current agrees very well with the observed value of 12.8 microampères. In this special case, where the concentrations of the thallous and potassium chlorides were equal, the current at the potential corresponding to the discharge of potassium ions was actually almost the same with or without thallous chloride present, and the apparent exaltation of the potassium ion limiting current can be regarded as due entirely to the exaltation of the migration current of the thallous ions themselves.

Although the foregoing exaltation phenomena have little practical value, they have important theoretical significance, because they furnish further confirmation of Heyrovský's theory of the migration current.

VIII. MAXIMA ON CURRENT-VOLTAGE CURVES

1. *The Heyrovský-Ilkovič theory of maxima*

One of the characteristic features of the current-voltage curves with the dropping mercury electrode is the more or less pronounced maxima, which are often present unless special measures are taken to prevent their occurrence. Maxima may also occur with quiet mercury electrodes (16, 26, 28), but these are not reproducible, in contradistinction to those obtained with the dropping mercury electrode, which are perfectly reproducible (48). The shapes of maxima vary from very acute peaks, with a rapid and almost discontinuous decrease of the current after the maximum, to rounded humps with a gradual decrease in current following the maximum (figure 15).

The shapes of current-voltage curves showing maxima are generally independent of the direction in which the applied E.M.F. is changed; if, after the maximum has been passed by increasing the E.M.F. in the usual way, the E.M.F. is then decreased gradually, the curve will in most cases retrace itself exactly over the maximum (figure 16). In certain cases, when the maxima are very acute, the potential at which the peak of the

maximum is reached may be somewhat different with decreasing E.M.F. from that with increasing E.M.F. For example, Heyrovský and Vas-

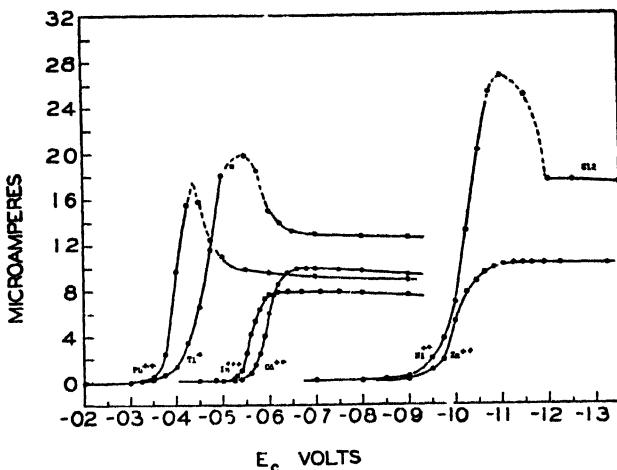


FIG. 15 Typical current-voltage curves of various metal ions in 0.1 N potassium chloride, showing various types of maxima.

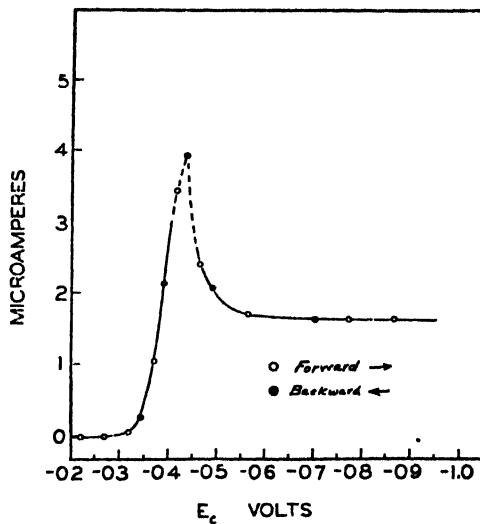


FIG. 16. Independence of the shape of the lead maximum of the direction in which the E.M.F. is applied. $2.75 \times 10^{-4} M$ lead nitrate in 0.1 N potassium chloride.

cautzau (48) found that the peak of the oxygen maximum was about 50 millivolts more negative with increasing applied E.M.F. than with

decreasing applied E.M.F., but the height of the maximum was practically the same in the two cases. We have verified this behavior.

Just as on other parts of a current-voltage curve, the current at the maximum appears to be independent of the time of electrolysis. If the applied E.M.F. is kept constant at a value at or near the peak of a maximum, the current remains constant indefinitely.

The maximum due to a given electroreducible substance may be either acute or rounded, depending on the composition of the solution. For example, Varasova (122) found that the acute maximum of oxygen (figure 13) was very small in extremely dilute potassium chloride solutions, but that it became more and more pronounced with increasing potassium chloride concentration and attained its greatest value in about 0.001 *N* potassium chloride. (Mr. C. S. Miller in this laboratory found the most pronounced maximum at a somewhat greater concentration of potassium chloride.) With increasing salt concentration the maximum decreased and became more rounded. Quite generally, a maximum will be most pronounced at a certain intermediate concentration of foreign salt; in distilled water and in concentrated salt solutions the oxygen maximum disappears (47).

An interesting feature when the maximum is very acute is that the current voltage curve often is a straight line, and not the usual S-shaped curve, before the peak of the maximum is reached. In such a case the slope of the straight line is proportional to the reciprocal of the resistance of the cell in accordance with Ohm's law (52). This indicates that the potential of the dropping electrode remains constant from the beginning of the discharge and that no concentration polarization occurs until the maximum is reached. Immediately after the maximum is passed and the current decreases suddenly, the dropping electrode becomes almost completely polarized. In weakly acid mercurous nitrate solutions this sudden increase in polarization may amount to as much as 1 volt (43). Conclusive proof that the electrode remains depolarized during the discharge of mercurous ions until the peak of the maximum is reached and that polarization sets in at the peak was obtained by Heyrovský from measurements of the interfacial tension at the mercury-solution interface during the electrolysis (35). He found that the interfacial tension, in a slightly acid mercurous nitrate solution, remained practically constant until the peak of the maximum was reached, instead of changing according to the electrocapillary curve. As soon as the maximum was passed, the interfacial tension changed abruptly and followed the electrocapillary curve.

It is still problematical whether the slope of the straight part of the current-voltage curve preceding an acute maximum corresponds exactly

to the reciprocal of the cell resistance. Ilkovič (54) assumes that the resistance of the surface film surrounding the electrode is different from that in the bulk of the solution, and that it varies with the distance from the electrode. In this connection reference is made to the important investigations of Hoekstra (49), who studied the current-voltage curves obtained with scraped electrodes. He found a constant apparent resistance at various values of the current, which was several times greater than that calculated from the dimensions of the cell and the specific conductance of the solutions (see also reference 12).

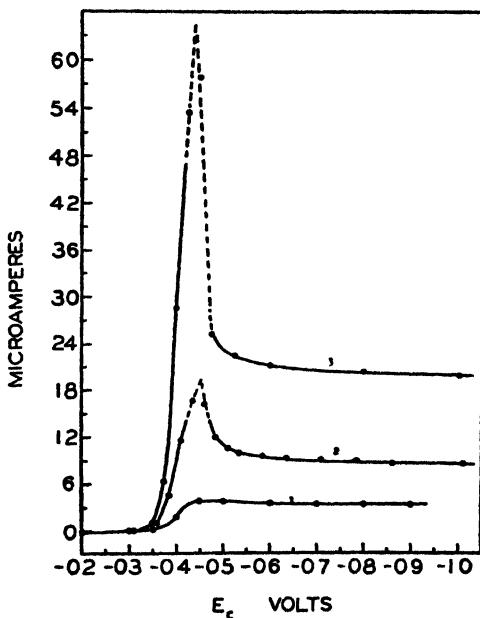


FIG. 17. Relative heights of maxima obtained with various concentrations of lead ions in 0.1 N potassium chloride. Curve 1, 3.3×10^{-4} M lead nitrate; curve 2, 1×10^{-5} M lead nitrate; curve 3, 2.3×10^{-5} M lead nitrate.

The height of the maximum of a given electroreducible substance is greatly dependent on its concentration. Reducible metal ions in dilute alkali chloride solutions usually do not yield maxima when the concentration of the metal ion is very small, but maxima appear and become more and more pronounced as the concentration of the metal ion is increased (figure 17). As a rule, there is no linear relation between the height of a maximum and the concentration of the reducible substance. For example, in the electrolysis of nickel chloride solutions, without any foreign salt present, Emelianova and Heyrovský (11) found that the ratio

i_{\max} / i_l was equal to 1:1 in a 0.0005 N solution, to 2.7:1 in a 0.0025 N solution, to 6:1 in a 0.0062 N solution, and to 6.5:1 in a 0.01 N solution.

The magnitudes of the maxima depend on the drop time, becoming smaller the slower the drop time (35).

The exact interpretation of the maxima is one of the most difficult problems in polarography. Heyrovský (35) attributes the maxima to an adsorption of the electroreducible substance on the growing mercury drops, whereby the concentration of the reducible substance is increased above its value in the body of the solution and normal concentration polarization is prevented. This adsorption is supposed, by Heyrovský, to be caused by the unhomogeneous electric field around the charged drops, which he assumes is identical with the electrokinetic potential at the mercury solution interface. Since pronounced maxima are found only with the dropping electrode, and since the adsorptive force must be established sooner than the electrokinetic potential can be built up, Ilkovič (54) assumes that the unhomogeneous field responsible for the adsorption is caused by the charging current (not the total current) rather than by the electrokinetic potential. The drop in potential around a mercury drop caused by the charging current, i_c , is $i_c W$, where W is the resistance in the surface film surrounding the drop. From fairly involved calculations Ilkovič concludes that the electric field created by the charging current is much greater than that due to the electrokinetic potential gradient. If the electric field due to the charging current were homogeneous, i.e., if $d(i_c W)/dx$ were constant and independent of the distance x from the surface of the electrode, it would cause only a motion of ions, and particles subject to dielectric polarization would simply become polarized and oriented, but would undergo no movement. However, the field created by the charging current is not homogeneous, because W varies greatly with x ; hence the unhomogeneous field causes a motion of dipolar molecules as well as ions. Ilkovič assumes that the attraction of dipolar molecules to the surface of the mercury drop by the unhomogeneous electric field takes place to an appreciable extent only when the concentration of the reducible substance has been reduced to a small value by the discharge process.

To clarify Ilkovič's view the following citation is given (54). "Let us now, from this viewpoint, explain the case of a maximum due to the depolarizing action of a solution which contains only one electrolyte, e.g., that due to the deposition of nickel from the solutions of its chloride. At voltages at which the rate of the deposition of nickel ions is smaller than the rate of their adsorption, the current is an 'adsorption' one and can increase with increasing voltage. However, as soon as the voltage is reached at which at the *very beginning* of the formation of the drop the

rate of deposition of the ions exceeds the rate of their adsorption, the exhaustion of the surface layer starts. At this instant a large electric force acts at the cathode, which cannot be depolarized, there being no reducible matter present in the exhausted surface film. However, the water molecules of the surface film are polarizable owing to their dipole character. They must be, therefore, polarized in the strong field, their oriented polarized layer thus constituting the increase of the 'back E.M.F.' observed at the fall of the maximum. Under such conditions now the electric field in the solution is considerably shortened so that no more ions can be attracted to the surface by adsorptive force. The current due to the deposition of nickel ions is now furnished only by the penetration of these cations through the polarized layer of water molecules and thus becomes a 'diffusion current.' The suppressive action of adsorbable particles on the maximum is similar to that of water molecules, *viz.*, in that they increase the polarization owing to their dipole moment and thus counterbalance the unhomogeneous electric field due to the charge of the small mercury electrode."

The Heyrovský Ilkovič interpretation of the maxima involves certain difficulties. We have seen in a previous section that both the charging current and the charge of the double layer (electrokinetic potential) change sign at the maximum in the electrocapillary curve. It is evident, therefore, that the *direction* of the electric field responsible for the adsorption must also change at the electrocapillary maximum, regardless of whether this field is due to the electrokinetic potential (Heyrovský) or to the charging current (Ilkovič). Hence reducible ions of a given sign should be adsorbed, by electrical attraction, only on one side or the other of the electrocapillary maximum, but not on both sides. On the negative side of the electrocapillary maximum the field is in the right direction to cause adsorption of positive ions, but on the positive side of the electrocapillary maximum the field is in the opposite direction, and hence positive ions should be repelled, not attracted and adsorbed. Actually it is found, in contradiction of the Heyrovský-Ilkovič theory, that the discharge of positive ions gives rise to maxima on *both* sides of the electrocapillary maximum, e.g., lead and nickel ions are discharged on opposite sides of the electrocapillary curve, but the current-voltage curves of both of these ions show prominent maxima (figure 15).

There is also a possibility that the maxima may be connected in some way with an electrostatic stirring effect at the surface of the mercury drops. Frumkin and Bruns (16), working with a quiet mercury cathode, showed experimentally that with relatively large current densities the surface of the mercury is in active motion. This is due to differences in the potential and the surface tension at various spots on the mercury surface. Conse-

quently, local currents flow between such spots and cause a stirring effect in the solution close to the surface of the electrode, as shown by Bruns, Frumkin, Joša, Vanjukova, and Zolotarewskaja (10), and by Stackelberg, Antweiler, and Kieselbach (113). Antweiler (4) has obtained experimental evidence that such a stirring effect actually exists with the dropping mercury electrode in cases where maxima occur. There is probably some connection between the charging current and this stirring effect. When the mercury is uncharged (isoelectric point), it is claimed that no maximum, and apparently no stirring effect, occurs.

At present it is not possible to give an exhaustive treatment of all the factors responsible for the occurrence of maxima. From a practical viewpoint, and also in connection with the further development of the theory, various observations made by Heyrovský and his school regarding the suppression and elimination of maxima are of great interest, and will now be discussed.

2. Suppression and elimination of maxima

Heyrovský (35) distinguishes between "positive" and "negative" maxima, according to whether a given maximum occurs on the positive or the negative side of the maximum in the electrocapillary curve. In the absence of capillary-active substances the maximum in the electrocapillary curve is at about -0.6 volt against the normal calomel electrode. In figure 15 the lead maximum is a positive maximum, while that of nickel is a negative maximum.

Maxima can be suppressed, and in most cases eliminated, by adding to the solution traces of certain capillary-active electrolytes and non-electrolytes, and various capillary-inactive ions and charged colloids. In this connection it should be realized that the capillary activity of a given substance, and hence its ability to suppress maxima, depends on the sign of the charge on the mercury (35).

Considering first the effect of capillary-inactive ions, there seems to be a close relation between the sign of the charge of the mercury on the one hand, and the signs and valences of the ions on the other. Thus, in the suppression of positive maxima, trivalent anions are usually much more effective than divalent, and the latter more so than univalent anions, while the valence of the cation is practically without effect. The converse is true for negative maxima, in which case the valence of the cation plays the predominant rôle. Heyrovský (35) states, "The dilutions in which cations of different valency cause the same degree of suppression of a negative maximum are in the same ratio as their power of precipitating (flocculating) a negative lyophobic colloid, say arsenic trisulfide. Indeed, the Hardy-Schulze rule has been found to hold strictly for the suppression

of maxima; thus the dilutions at which ions of potassium, calcium, and lanthanum, respectively, produce suppression to the same extent were found to be in the ratio 1:160:10000." However, it is doubtful whether the Schulze-Hardy rule holds strictly in the suppression of maxima. For example, Emelianova and Heyrovský (11) found the negative maximum of nickel just suppressed in 0.11 *N* potassium chloride, in 2.2×10^{-3} *N* magnesium chloride, in 1.1×10^{-3} *N* calcium chloride, in 3×10^{-5} *N* barium chloride, in 2×10^{-6} *N* aluminum chloride, and in 1×10^{-6} *N* lanthanum chloride. It is not clear why barium ions in this respect behave as trivalent cations, quite differently from calcium and magnesium ions.

Mr. C. S. Miller in this laboratory found the difference in the suppressive effect of barium ions on the one hand, and of calcium and magnesium ions on the other, much smaller. He also studied the effect on the maxima of additions of a deficiency of sodium hydroxide to the nickel chloride solution. The alkali causes the partial separation of basic nickel salt, presumably as a positively charged colloid. One would expect, therefore, that the negative nickel maximum would be suppressed completely in the presence of this colloid. Actually it was found that the ratio of the maximum current to the limiting current remained almost unchanged in the presence of various amounts of the positive colloid. Another exception to the rule given by Heyrovský was found by Mr. Miller in the case of the positive oxygen maximum. In 0.001 *N* potassium chloride a very pronounced, acute maximum occurs, the maximum current being much greater than the limiting current (figure 13). Addition of potassium ferrocyanide to this solution, in concentrations ranging between 10^{-5} and 6×10^{-6} molar, hardly affected the maximum, and at a concentration of 0.001 *M* ferrocyanide the maximum was even slightly more pronounced instead of suppressed. Also at other concentrations of potassium chloride no pronounced effect of ferrocyanide ions was found. It is evident, therefore, that care must be exercised in generalizing Heyrovský's rule.

From Herasymenko's study (26) of the effect of salts on the maximum in the reduction of uranyl ions (first wave), cations of various valency (K^+ , Na^+ , Mg^{++} , Al^{+++}) were found to have the same effect, and there was also hardly any difference between the effect of chloride and sulfate ions. Since the uranyl ions (first wave) are reduced on the positive side of the electrocapillary maximum, one would expect that sulfate ions would be much more effective in suppressing the maximum than chloride ions.

The following peculiar phenomenon is also hard to explain on the basis of Heyrovský's rule. Heyrovský and Dillinger (43) found that in a mixture of nickel and manganous chloride two pronounced negative maxima occurred. The nickel maximum was suppressed by addition of

barium ions, but that of manganese was much less affected. On the basis of Heyrovský's rule both maxima should have been equally suppressed.

From a practical point of view the effect of *capillary-active* ions on maxima is of much greater importance than that of capillary-inactive ions. Heyrovský (38, 35) states that negative dyes and negative colloids (he used a negative colloidal solution of a barium soap) suppress the positive thallium maximum very easily, whereas positive dyes and positively charged colloids cause a suppression of negative maxima. For example,

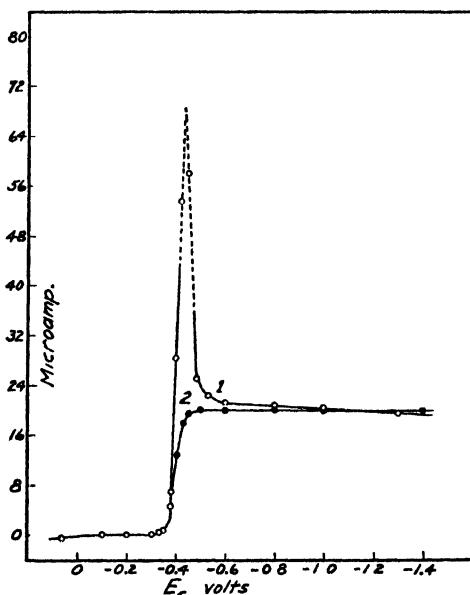


FIG. 18. Suppression of the positive lead maximum by anionic methyl red in neutral solution. Curve 1, 50 cc. of $2.3 \times 10^{-3} M$ lead nitrate in 0.1 *N* potassium chloride; curve 2, 0.1 cc. of 0.1 per cent sodium methyl red added.

the positive maximum of thallium is easily suppressed by acid fuchsin (anion), but that of nickel is hardly affected. On the other hand, basic fuchsin (cation) suppresses the negative nickel maximum completely.

In practical analytical work the occurrence of maxima is a nuisance, but fortunately they can usually be eliminated by the addition of suitable capillary-active ions. Lingane (65) found, for example, that the positive maxima of thallium and lead in neutral solutions are completely suppressed by traces of the sodium salt of the ordinary indicator methyl red (figure 18). However, in accordance with Heyrovský's rule, the positive lead maximum is not suppressed by the cation form of methyl red in acid

medium (figure 19). On the other hand, the negative maximum of nickel in neutral medium was found to be unaffected by the anion form of methyl red, but in acid medium the maximum was completely eliminated by the cation form of the dye; this is also in accord with Heyrovský's rule.

Methyl red is itself reduced at the dropping electrode, but the concentrations of the dye used to suppress maxima are so small (less than 0.001 per cent) that they hardly affect the diffusion current of the substance to be determined.

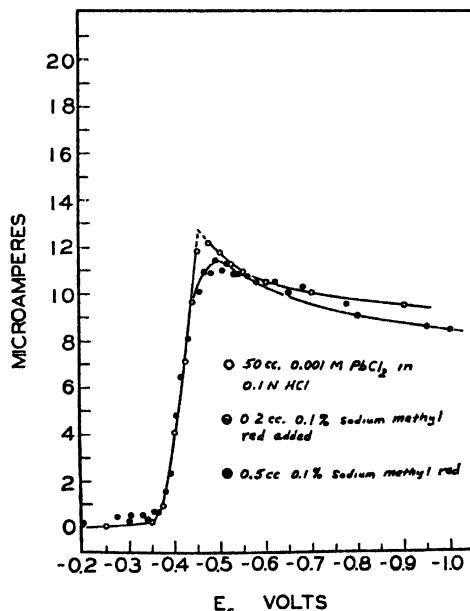


FIG. 19. Failure of cationic methyl red to suppress the lead maximum in acid solution.

The analytical significance of the elimination of maxima is demonstrated in figure 20 by the current-voltage curve of a mixture of lead and cadmium ions without and with methyl red present. In the absence of the dye no definite measure of the diffusion current of the lead ions can be obtained, but a well-defined diffusion current for lead results when the maximum is eliminated by the addition of methyl red (65).

One should not infer from the foregoing examples that strongly capillary-active anions invariably suppress only positive maxima, and capillary-active cations suppress only negative maxima. We have already cited several exceptions to this rule from the work of Heyrovský and Dillinger. The degree of polarizability of the ions, their specific ability to be adsorbed

on mercury at various potentials, and their effect of shifting the electrocapillary zero also have to be taken into account. That positive dyes can be as effective as negative dyes in suppressing the positive oxygen maximum is evident from the work of Rayman (98). Furthermore, Hamamoto (20) has demonstrated that positively charged alkaloid ions have a great suppressive effect on the positive oxygen maximum.

Capillary-active non-electrolytes can also be very effective in suppressing and eliminating maxima in certain cases. In general, it is not to be ex-

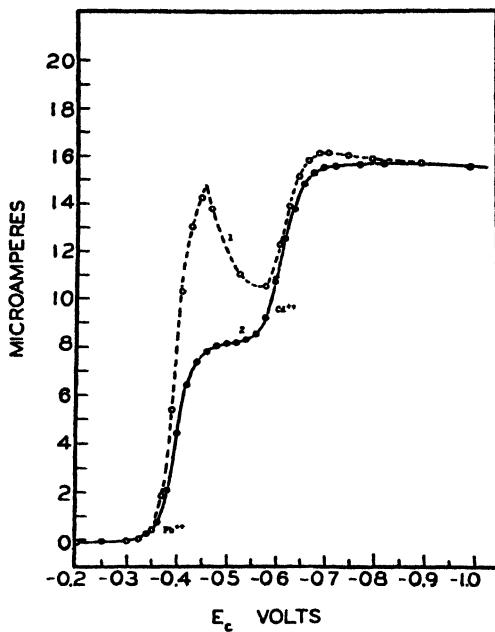


FIG. 20. Current-voltage curves of a neutral 0.1 *N* solution of potassium chloride containing 0.001 *M* lead and cadmium ions. Curve 1, no methyl red present; curve 2, 0.1 cc. of 0.1 per cent sodium methyl red added per 50 cc.

pected that a given capillary-active non-electrolyte will be able to suppress maxima over the entire voltage range, because the adsorbatibility of such substances usually reaches a maximal value at a certain characteristic potential. It is to be expected that sulfur-containing organic substances should be effective in eliminating maxima, since such substances are usually capillary-active. Heyrovský (35) mentions the great effect of traces of α -naphthol in suppressing maxima, and Maas (68) found that naphthalene was effective. Oxygen, which in its own electroreduction gives a very pronounced maximum, suppresses the maxima of various metal ions, e.g., nickel ions.

We (65) found that gelatin, even in very low concentrations, was able to suppress the maxima of various metal ions over a wide range of potentials. The use of gelatin to suppress maxima has also been recommended by Hohn (50). The suppressive effect of gelatin on the lead maximum is shown in figure 21. It will be noted that a concentration of gelatin of 0.02 per cent was ample to suppress the lead maximum completely. When the concentration of the gelatin is greater than about 0.01 per cent it decreases the diffusion currents, probably because it decreases the

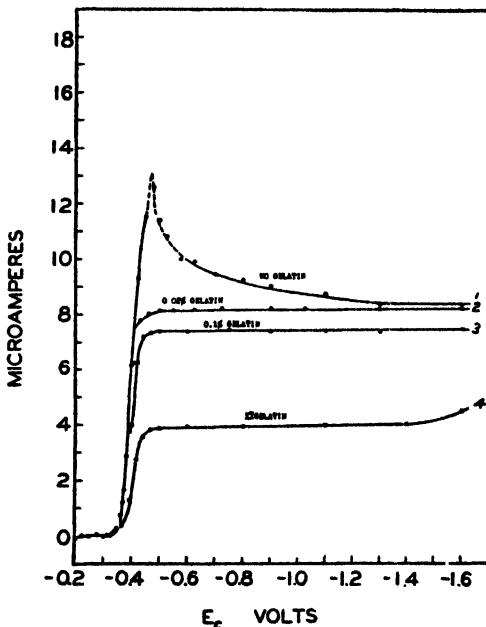


FIG. 21. Suppression of the lead maximum by gelatin in neutral potassium chloride solution, 0.001 M lead chloride in 0.1 N potassium chloride. Curve 1, no gelatin present; curve 2, 0.02 per cent gelatin; curve 3, 0.1 per cent gelatin; curve 4, 1 per cent gelatin.

diffusion coefficient of the reducible ions. We have also found that gelatin completely eliminates the maximum of thallium in both acid and alkaline medium, and that it eliminates the maximum of nickel in both neutral and ammoniacal medium. From the cases so far studied it appears that gelatin is effective in suppressing many maxima. In practical work, however, its suppressive effect on the diffusion current, when it is present at concentrations greater than about 0.01 per cent, must be taken into account. Furthermore, it is our experience that the addition of gelatin may actually be harmful, instead of beneficial, in

certain cases, particularly when the electroreduction is irreversible. A more extensive study of the influence of gelatin is necessary before it can be generally applied in practical analytical work.

3. *Phenomena at the electrocapillary zero*

Heyrovský claims that no maximum occurs when the electroreduction of a substance takes place at the potential corresponding to the electrocapillary zero (maximum in the electrocapillary curve), where the mercury is apparently uncharged and the interfacial tension has its maximal value (35). This statement may be correct so far as the electroreduction of capillary-inactive ions and molecules is concerned, but, to say the least, it is doubtful that it is true in the electroreduction of capillary-active substances (128). For example, no maximum occurs in the discharge of the capillary-inactive cadmium ions from chloride solutions, because the discharge potential of cadmium ions under these conditions is very close to the electrocapillary zero (-0.6 volt in chloride solutions) (figure 15). Cadmium iodide, on the other hand, gives a very pronounced maximum, even though the half-wave potential of cadmium in very dilute iodide solutions is practically the same as in chloride solutions. The occurrence of the (positive) maximum is explained (35) by the fact that the iodide ions are strongly capillary-active, and they shift the electrocapillary zero to a more negative potential so that it no longer coincides with the discharge potential of the cadmium ions. Other capillary-active anions, such as S^{--} , CN^- , CNS^- , etc., shift the electrocapillary zero in a similar way, while capillary-active cations shift it to a more positive potential (15).

Care must be exercised in the interpretation of the effect of these capillary-active ions on maxima. Capillary-active anions are adsorbed on positively charged and uncharged mercury (increasing adsorption with increasing positive charge), and they therefore attract an almost equivalent amount of cations to the surface in a triple layer (123). For this reason the total effect of capillary-active ions may be very complicated. For example, Lingane (65) found that the positive maximum due to lead ions in 0.1 N potassium chloride was almost completely suppressed when the solution was made 0.001 N with respect to iodide ions. On the other hand, in the absence of potassium chloride the lead maximum was considerably accentuated in 0.001 N and 0.005 N potassium iodide, and was not completely eliminated even in 1 N potassium iodide (in the latter case the lead was present chiefly as PbI_4^- anions). It is probable that a similar effect of iodide ions will be found in the case of cadmium and other metal ions.

Heyrovský and Vascautzanu (48) have demonstrated the effect of

thiocyanate ions in shifting the electrocapillary zero and the resultant effect on the slight negative maximum of cadmium nitrate. They found that the slight negative maximum in a 0.005 *M* cadmium nitrate solution was eliminated by the addition of sufficient potassium thiocyanate to make its concentration 0.001 *N*. With further additions of thiocyanate ions the electrocapillary zero was shifted to more negative potentials, and the cadmium maximum reappeared on the positive side of the electrocapillary zero.

Similar phenomena were observed (48) in the electroreduction of (undissociated) maleic acid. The current-voltage curve of 0.02 *M* maleic acid in 1 *N* hydrochloric acid shows a pronounced maximum (note the occurrence of a maximum at this high electrolyte concentration), which, according to Heyrovský and Vascautzanu, is a negative maximum. When the maleic acid solution was made about 0.002 *N* with respect to potassium iodide the maximum disappeared, but it reappeared again as a positive maximum when the electrocapillary zero was shifted to still more negative potentials by further addition of iodide.

Instead of changing the location of the electrocapillary zero, use can be made of complex formation to shift the discharge potentials of metal ions which are normally reduced on the positive side of the electrocapillary zero to more negative values, until they coincide with the electrocapillary zero, and the maxima then disappear. By adding a large excess of the complex-forming salt, the discharge potentials can, in certain cases, be shifted to such negative values that the maxima are changed from positive to negative maxima (48).

4. Influence of external resistance in the cell circuit on the maxima

Brdička (7) has recently shown that the acute maximum accompanying the discharge of mercurous ions from a slightly acid mercurous nitrate solution was greatly suppressed and almost completely eliminated by a high resistance in series with the cell. As the external resistance was increased from 26 to 5000 ohms the maximum decreased markedly, but the peak of the maximum was at the same value of the total applied E.M.F., that is, at the same value of the *total* potential drop across both the external resistance and the cell. The current-voltage curves during the deposition of mercury (depolarized electrode) were straight lines up to the peak of the maximum, the slope being determined by the total resistance according to Ohm's law. The maximum occurred at the same total applied E.M.F., independent (within a certain range) of the external resistance, which means that the potential of the dropping electrode itself at the peak of the maximum was shifted to more positive values with increasing external resistance, or that the value of E_e at the maximum

became smaller with increasing external resistance. At first glance this effect of external resistance is difficult to understand, since with the same value of E , we would expect the same current, independent of the external resistance. This actually is the case with the quiet mercury electrode, since in this case the external resistance cannot have an effect upon the height and location of the maximum.

In order to interpret the peculiar effect obtained with the dropping electrode, we must realize that the measured values of the current and of E_{cell} are average values; the actual values of i and E_{cell} vary during the growth of the drop. In the one extreme case in which the external resistance is zero, the applied E.M.F. is equal to E_{cell} . Under these conditions the latter is constant during the growth of the drop, and, as will be shown below, the current during the formation of a drop when the electrode is *depolarized* will increase with the area of the drop. In the other extreme case, in which the internal resistance is negligibly small with regard to the external resistance, E_{cell} will increase during the growth of the drop when the electrode is depolarized, but the current will remain practically constant after the very early stage of the formation of the drop. Hence the fluctuation of the galvanometer when dealing with a depolarized dropping electrode should decrease with increasing external resistance and finally should disappear.

Let us now compare the change of the current during the formation of a drop (depolarized electrode) in the two extreme cases when the average E_{cell} is the same, but when the external resistance in one case is equal to zero and in the other case is infinitely great. Naturally, with the same average E_{cell} the average current measured will be nearly the same. In the absence of an external resistance the current should increase during the formation of the drop, but, on the other hand, with a large external resistance in the circuit the current remains practically constant after the very early stages of the formation of the drop. As the average current is the same in both cases, the *current density* during the early stages of the formation of the drop in the presence of a large external resistance is much greater than in the absence of external resistance. It has been stated earlier in this section that an acute maximum is characterized by a sudden polarization due to an exhaustion of depolarizer in the immediate vicinity of the growing drop. The drop then remains polarized, and the current drops from the maximum to the value of the diffusion current. Since the current density during the early stages of the formation of the drop is, with the same average E_{cell} , much greater with a large external resistance in the circuit than in the absence of an external resistance, the critical exhaustion of the depolarizer (maximum) will be obtained at a smaller average E_{cell} when the external resistance is great than when there is no external resistance in the circuit. From this it follows that the

height of the maximum must also decrease with increasing external resistance. Brdička (7) calculated the change of the current with time as a function of the external resistance. His derivations are given below in a condensed form.

When the dropping electrode is depolarized and the external resistance is equal to zero, the current is determined by Ohm's law:

$$i = \frac{E_a}{W_i} \quad (42)$$

in which W_i is the internal resistance and E_a is equal to the total applied E.M.F. The internal resistance changes with the size of the mercury drop. According to Ilkovič (54)

$$W_i = \frac{\rho}{4\pi r} \quad (43)$$

in which ρ denotes the specific resistance of the electrolyte and r the radius of the mercury drop considered to be spherical. Earlier in this paper it has been shown that r is a simple function of the time t ,

$$r = at^{1/3} \quad (44)$$

in which a is a function of the characteristics of the capillary used (equation 6). From equations 42, 43, and 44 it follows that

$$i = \frac{E_a 4\pi a t^{1/3}}{\rho} \quad (45)$$

The average current, i' , according to Ilkovič (52) is

$$i' = \frac{4\pi E_a a}{t_{\max.}} \int_0^{t_{\max.}} t^{1/3} dt = \frac{E_a}{W'} = \frac{E_a}{4/3 W_{\min.}} \quad (46)$$

in which W' is the mean resistance which is equal to $4/3 W_{\min.}$, where $W_{\min.}$ is the minimum resistance at the moment when the drop falls. When the drop is depolarized and the external resistance in the circuit is equal to W_e , the current at any time t is given by

$$i = \frac{E_a}{W_i + W_e} = E_a \frac{4\pi a t^{1/3}}{\rho + 4\pi a t^{1/3} W_e} \quad (47)$$

and the mean current, i' , by

$$i' = \frac{E_a 4\pi a}{t_{\max.}} \int_0^{t_{\max.}} \frac{t^{1/3} dt}{\rho + 4\pi a t^{1/3} W_e} \\ = \frac{E_a}{W_e} \left\{ 1 - \frac{3}{2} \frac{W_{\min.}}{W_e} + 3 \left(\frac{W_{\min.}}{W_e} \right)^2 - 3 \left(\frac{W_{\min.}}{W_e} \right)^3 \ln \left(1 + \frac{W_e}{W_{\min.}} \right) \right\} \quad (48)$$

The mean resistance, W' , is given by the expression,

$$W' = \frac{W_*}{2} + \frac{3}{2} \frac{W_{\min}}{W_*} + 3 \left(\frac{W_{\min}}{W_*} \right)^2 - 3 \left(\frac{W_{\min}}{W_*} \right)^3 \ln \left(1 + \frac{W_*}{W_{\min}} \right) \quad (49)$$

Brdička writes this expression in a simplified form:

$$W' = W_* + \alpha W_{\min} \quad (50)$$

in which α depends upon the ratio W_{\min}/W_* as shown in the following table:

| W_{\min}/W_* | 1/1000 | 1/100 | 1/10 | 1 | 10 | 100 |
|----------------|--------|-------|-------|-------|-------|-------|
| α | 1 499 | 1 493 | 1.457 | 1 378 | 1 336 | 1 334 |

In the electrolysis of a slightly acid mercurous nitrate solution, the minimum resistance of which was 1008 ohms, Brdička found that equation 50 was valid within 2 per cent when the external resistance was varied from 26 to 8026 ohms.

From equation 45 it is evident that with the depolarized electrode the current should increase with $t^{1/3}$ during the formation of a drop when the external resistance is zero. On the other hand, it is seen from equation 47 that the current should not change during the formation of the drop when the external resistance becomes so great that ρ is negligibly small compared to $4\pi a t^{1/3} W_*$. If this extreme condition is fulfilled, we can write, instead of equation 47

$$i = \frac{E_a}{W_*}$$

Brdička (7) determined current time curves during the formation of a drop in the neighborhood of the maximum in slightly acid mercurous nitrate solutions with the aid of a galvanometer of extremely short period with various external resistances in the circuit. In the absence of an external resistance the current increased continuously and rapidly during the formation of each drop, apparently in direct proportion to the increase of the area. With increasing external resistance the rate of increase of the current during the growth of the drop became increasingly less. Indeed, with a large external resistance the current became almost constant after the early stages of drop formation. The current-time curves obtained by Ilković (1936) in the electrolysis of oxygen in 0.002 N potassium chloride solutions in the immediate neighborhood of the maximum are also of interest in this connection and can be interpreted in a similar way.

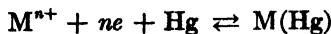
IX. EQUATIONS OF POLAROGRAPHIC WAVES AND THE SIGNIFICANCE OF THE HALF-WAVE POTENTIAL

The steeply rising part of a current-voltage curve between the decomposition potential and the potential at which a constant limiting or diffusion current is reached is known as a "polarographic wave." In this section we shall consider the relations that exist between the potential of the dropping electrode and the corresponding current at each point on a polarographic wave, that is, the equations of polarographic waves.

In the usual type of electrolysis experiment, employing large electrodes, relatively large concentrations of reducible or oxidizable substances, and thorough stirring, the concentration polarization is reduced to a minimum and the current increases almost without limit after the decomposition potential is exceeded. Under these conditions the current-voltage curve is practically a straight line, which follows Ohm's law, with a slope equal to the reciprocal of the resistance of the cell. On the other hand, in electrolyses with the dropping mercury electrode, and platinum micro-electrodes in unstirred solutions, conditions are such that concentration polarization is favored as much as possible, rather than eliminated. Hence the normal polarographic waves (no maximum present) obtained with the dropping electrode are not straight lines but definitely S-shaped curves. The current at each point on a normal polarographic wave is an exponential function of the potential of the dropping electrode.

1. Electrodeposition of simple metal ions

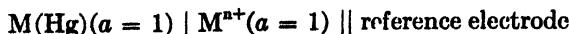
If the electrodeposition of metal ions at the dropping mercury cathode takes place reversibly according to the equation



then the potential at each point on the polarographic wave should be given by

$$E = E_a^0 - (RT/nF) \ln (a_e/a_0) \quad (51)$$

In this equation a_e is the activity of the deposited metal atoms in the amalgam formed on the surface of the mercury drops, a_0 is the activity of the reducible metal ions in the layer of solution at the *surface of the drops*, and E_a^0 is the standard or normal potential of the metal amalgam that is, the E.M.F. of the cell



If we introduce the usual relation between concentration and activity, equation 51 becomes

$$E = E_a^0 - (RT/nF) \ln (f_e/f_0) - (RT/nF) \ln (C_e/C_0) \quad (52)$$

in which f_a is the activity coefficient of the metal atoms in the amalgam and f_s is the activity coefficient of the reducible metal ions in the solution.

When an excess of foreign salt is present in the solution, the reducible ions reach the surface of the dropping electrode entirely by diffusion, and the current is governed by the rate of diffusion. In turn, the latter will depend directly on the difference in concentration between the depleted surface layer and the body of the solution, so Heyrovský and Ilkovič (44) write

$$i = K(C - C_0)$$

or

$$C_0 = C - (i/K) \quad (53)$$

in which i and C_0 are the current and the corresponding concentration in the depleted surface layer at each point on the wave, and C is the concentration in the body of the solution. When a constant diffusion current is reached, C_0 will be reduced to a constant minimum value, and since this minimum value is negligibly small compared to C , we have

$$i_d = KC \quad \text{or} \quad C = i_d/K \quad (54)$$

The validity of this equation, in which K is the diffusion current constant, has already been demonstrated.

Heyrovský and Ilkovič (44) assume that the average concentration of the amalgam at any point on the wave is directly proportional to the current, that is, $C_a = ki$. According to these authors (44) k is inversely proportional to the square root of the diffusion coefficient of the metal atoms in mercury.

When the foregoing relations are introduced into equation 52, we obtain

$$E = E_a^0 - (RT/nF) \ln (f_a k K / f_s) - (RT/nF) \ln (i/(i_d - i)) \quad (55)$$

where E and i are corresponding values at any point on the polarographic wave.

Now the half-wave potential, designated by $E_{1/2}$, is defined as the potential of the dropping electrode at the point on the polarographic wave where $i = i_d/2$. It is evident that when $i = i_d/2$ the last log term in equation 55 is zero, and hence we find that

$$E_{1/2} = E_a^0 - (RT/nF) \ln (f_a k K / f_s) \quad (56)$$

and

$$E = E_{1/2} - (RT/nF) \ln (i/(i_d - i)) \quad (57)$$

This is the equation of the polarographic wave in the electrodeposition of simple metal ions, first derived by Heyrovský and Ilkovič (44).

If equation 57 is valid, it is evident that a plot of E against $\ln(i/(i_d - i))$ should give a straight line with a slope equal to RT/nF , and the potential where the log term is equal to zero should be the half-wave potential. This prediction and the general validity of equation 57 in the deposition of various metal ions have been verified experimentally by Tomeš (119) and Lingane (65). For example, the current-voltage curve of 0.001 M thallous chloride in 0.9 N potassium chloride, together with the straight line obtained by plotting E against $\log(i/(i_d - i))$, is given in figure 22. The slope of the straight line is 0.061 volt, in good agreement with the

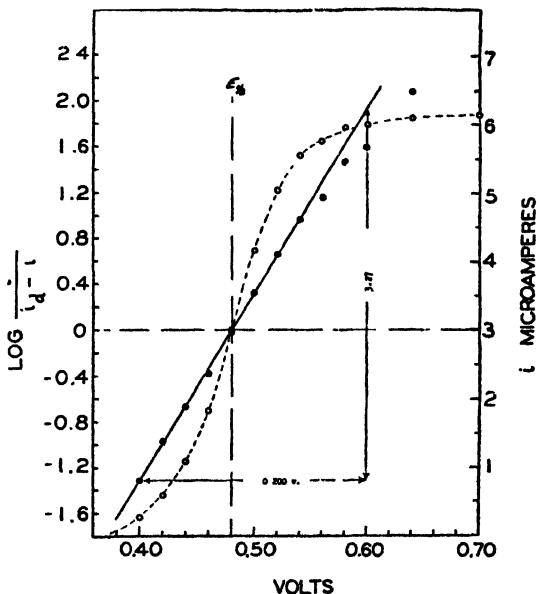


FIG. 22. Test of equation 57 in the discharge of 0.001 M thallous chloride from 0.9 N potassium chloride.

theoretical value 0.059 volt for the electrodeposition of a univalent metal ion, and the potential corresponding to a value of zero for the log term is the half-wave potential of the current-voltage curve. Tomeš (119) and Lingane (65) have also verified equation 57 in the electrodeposition of lead, cadmium, and indium ions from 0.1 N and 1 N potassium chloride. Some typical results (65) are shown in figures 23 and 24. In all these cases the log plot is a good straight line, with a slope in close agreement with the theory, which should be 0.030 volt for cadmium and lead ions and 0.020 volt for trivalent indium ions.

The foregoing derivations are based on the assumption that the rate of

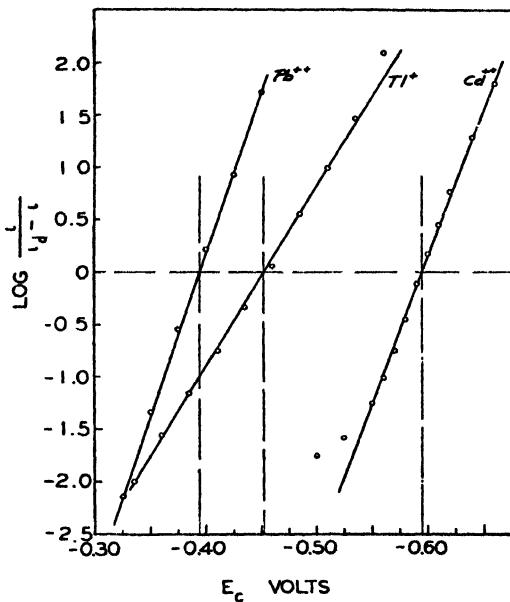


FIG. 23. Test of equation 57 in the discharge of lead, thallous, and cadmium ions from 0.1 N potassium chloride.

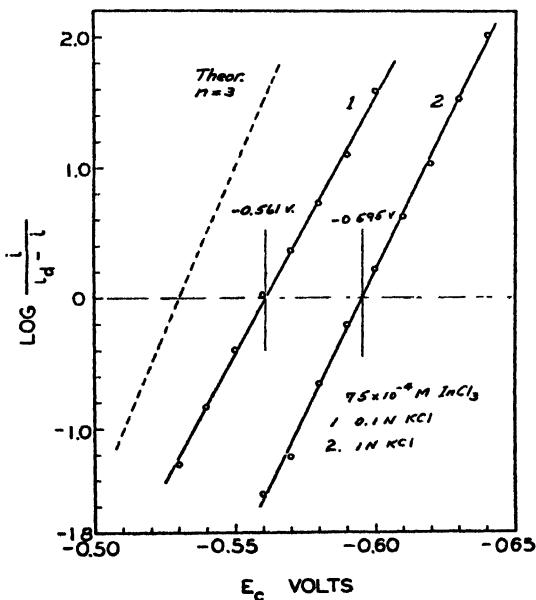


FIG. 24. Test of equation 57 in the discharge of indium ions from 0.1 and 1 N potassium chloride.

diffusion is the current-determining factor. If this condition is not fulfilled, the slope of the current-voltage curve will be less than corresponds to equation 57. In such a case the plot of E against $\ln(i/(i_d - i))$ may still be a straight line, but the slope may not correspond to the true value of n . For example, we (65) found that the log plot of the current-voltage curve of nickel in ammoniacal medium was a straight line, but that its slope was less than corresponds to the known n value of 2. When gelatin was added to the solution, the slope deviated still more from the theoretical value. This abnormal behavior may be caused by a certain degree of irreversibility ("chemical polarization") in the deposition of nickel ions under these conditions. We (65) found that the slope of the log plot of the current-voltage curve of lead in neutral medium is also markedly changed (apparent value of n less than 2) when gelatin is present in the solution.

By differentiating equation 57 we obtain

$$\frac{dE}{di} = - (RT/nF)[(1/i) + (1/(i_d - i))] \quad (58)$$

$$\frac{d^2E}{di^2} = (RT/nF)[(1/i^2) - (1/(i_d - i)^2)] \quad (59)$$

Since the slope of the polarographic wave is di/dE , we see from equation 58 that the slope or steepness of the wave depends on both the valence, n , and the concentration of the reducible ions. Hence at a given value of E the slope of the wave should be greater, the greater the concentration of the reducible ions, and at a given concentration the steepness of the wave should be greater, the greater the valence of the reducible ions. In other words, at a given concentration the wave of a divalent ion should rise more steeply than that of a univalent ion. These characteristics of the waves are actually found experimentally.

By setting the second differential (equation 59) equal to zero, we find that the curve of equation 57, and hence the polarographic wave, should have a point of inflection when i is equal to $i_d/2$. The half-wave potential thus corresponds to the true inflection point of the wave, and is therefore a significant constant for each reducible metal ion (44).

We see from equation 56 that $E_{1/2}$ should depend on the values of E_a^0 , K , k , f_a , and f_s . At a given concentration of foreign salts, E_a^0 , K , and f_s will be constant and independent of the concentration of the reducible ions. Since the amalgams formed on the surface of the mercury drops are usually very dilute, we may safely assume that they behave as ideal solutions, and hence it is to be expected that k and f_a will also be almost constant in the range of amalgam concentrations involved in polarographic electrolyses. It is to be expected, therefore, that with a given composition of the solution with respect to foreign salts, the values of $E_{1/2}$ should be

practically constant and independent of the concentration of the reducible ions (38, 44). This has actually been found to be true by many investigators, and it is demonstrated by the data in table 8 that we have obtained (65). The constancy of $E_{1/2}$ at various concentrations of the reducible ions is obviously one of the most important characteristics of polarographic waves from the standpoint of qualitative polarographic analysis.

Since f_s will decrease with increasing ionic strength, it is evident from equation 56 that the half-wave potentials of metal ions should be shifted to somewhat more negative values as the concentration of foreign salt is increased. However, since only the logarithm of f_s is involved, a considerable change in the activity coefficient should have only a relatively small effect on the half-wave potential.

TABLE 8

Constancy of the half-wave potentials of several metal ions at various concentrations in 0.1 N potassium chloride

Half-wave potentials at 25°C. in volts, referred to the saturated calomel electrode

| C millimoles per liter | HALF-WAVE POTENTIALS | | | |
|---------------------------|--------------------------|---------------------------|---------------------------|---------------------------|
| | Tl ⁺ volts | Pb ⁺⁺ volts | Cd ⁺⁺ volts | Zn ⁺⁺ volts |
| 0.05 | (-0.45) | | | |
| 0.1 | -0.462 | -0.396 | | |
| 0.2 | | | -0.594 | -0.990 |
| 0.5 | 0.457 | 0.396 | 0.593 | 0.989 |
| 1.0 | 0.460 | 0.392 | 0.594 | |
| 2.0 | 0.456 | 0.397 | 0.601 | 0.999 |
| 5.0 | 0.459 | 0.394 | 0.598 | 0.992 |
| 10.0 | | 0.398 | 0.605 | -1.003 |
| Average... | -0.459 ± 3 | -0.396 ± 3 | -0.599 ± 5 | -0.995 ± 6 |

It should also be mentioned that the half-wave potential of a given metal ion is usually not influenced by the preceding discharge of other metal ions or uncharged substances. The half-wave potentials of metal ions in mixtures are, in general, the same as in their single solutions. In fact, the same is true for most reducible substances, both inorganic and organic. The half-wave potential is also independent of the particular capillary used and of the rate of dropping (44).

According to Heyrovský and Ilkovič (44), the product kK in equation 56 is equal to $(D_s/D_a)^{1/2}$, where D_a is the diffusion coefficient of the metal atoms in the amalgam formed on the surface of the mercury drops and D_s is the diffusion coefficient of the reducible metal ions in the solution. Heyrovský and Ilkovič assume that the quantity $(D_s/D_a)^{1/2}$ is practically

equal to unity. On the basis of this assumption, and since the quotient f_a/f_s will also be of the order of unity, it follows that $E_{1/2}$ for a given metal ion should be at least approximately equal to E_a^0 . Furthermore, if we employ the usual convention that the activity of a solid metal is unity, then E_a^0 will be identical with the ordinary standard or normal potential of the metal, E_M^0 . Hence the half-wave potentials of the various metal ions should be about the same as the ordinary standard potentials. This is demonstrated by the following comparison of the values of $E_{1/2}$ and E_M^0 of various metals, referred to the saturated calomel electrode. The $E_{1/2}$ values of copper, sodium, and potassium ions have been taken from the compilation given by Heyrovský (38)

| | Cu^{++} | Pb^{++} | Ti^+ | Cd^{++} | Zn^{++} | Na^+ | K^+ |
|--------------------|------------------|------------------|---------------|------------------|------------------|---------------|--------------|
| $E_{1/2}$ vs S C E | +0 12 | -0 396 | -0 459 | -0 599 | -0 905 | -2 12 | -2 14 |
| E_M^0 vs S.C.E | +0 10 | -0 372 | -0 582 | -0 647 | -1 008 | -2 96 | -3 17 |

It will be noted that the $E_{1/2}$ values are in most cases fairly close to the ordinary reversible standard potentials, except for the alkali metal ions whose half-wave potentials are of the order of 1 volt more positive than the corresponding standard potentials.

The experimentally determined half-wave potentials of a large number of inorganic ions are shown in figure 25. This chart was compiled by V. Majer (44), and it is appropriately called a "polarographic spectrum" in analogy to the ordinary optical spectrum of the elements.

In certain cases the reduction of metal ions at the dropping electrode may proceed simply to a lower valence state rather than to the metallic state. This is the case, for example, in the reduction of ferric, chromic, and cobaltic ions, which are respectively reduced to ferrous, chromous, and cobaltous ions. In such a case the reduction may be represented by



If this reduction takes place reversibly, the potential of the dropping electrode at any point on the wave should be given by

$$E = E^0 - (RT/mF) \ln (f_{\text{red}}/f_{\text{ox}}) - (RT/mF) \ln (C_{\text{red}}^0/C_{\text{ox}}^0) \quad (60)$$

where C_{red}^0 and C_{ox}^0 are, respectively, the concentrations of the reduced and oxidized forms at the surface of the dropping electrode, and f_{red} and f_{ox} are the corresponding activity coefficients. It is evident that in this case E^0 is the ordinary standard oxidation-reduction potential of the system. In such a reduction the dropping electrode simply functions as an indifferent electrode. If the solution originally contains none of the reduced form of the metal ions, it may be readily shown that the equation of the

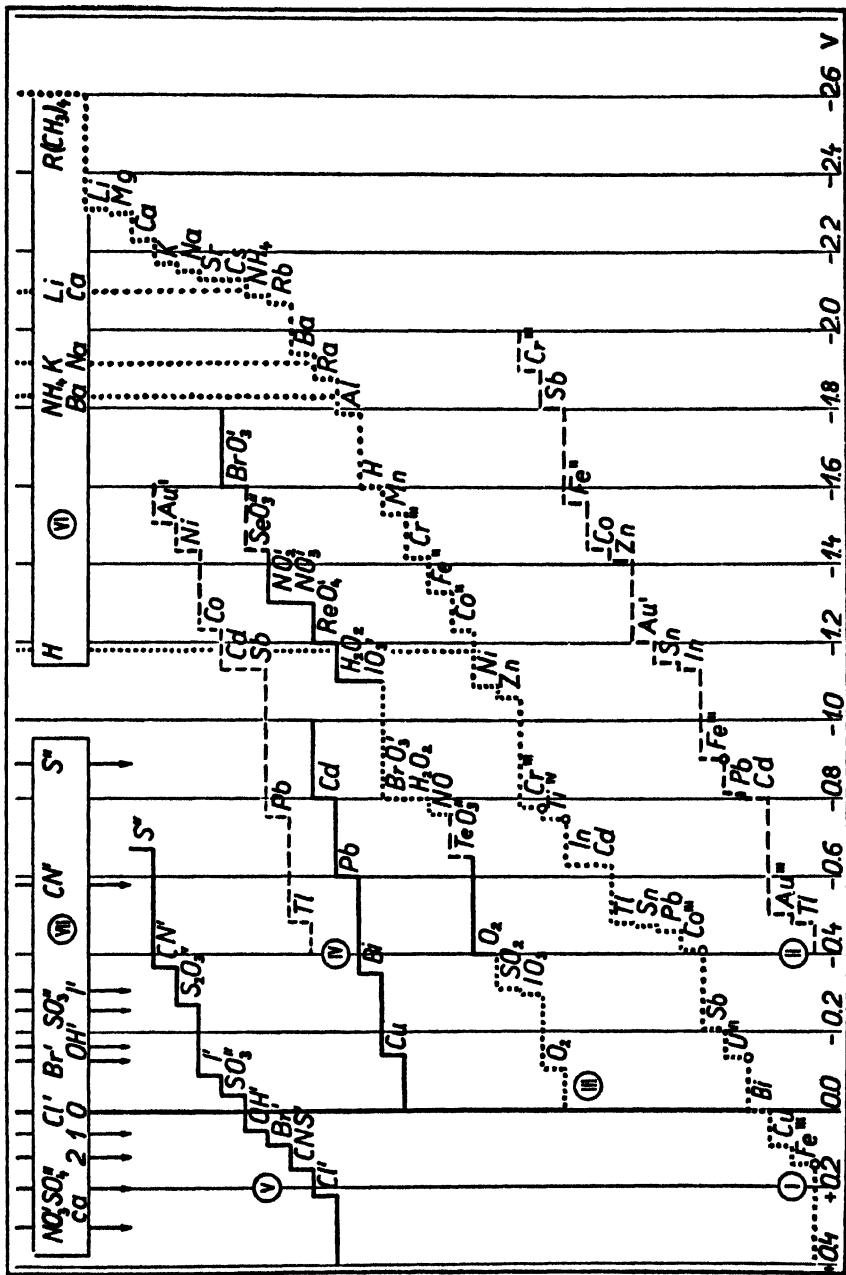


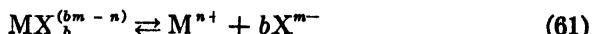
FIG. 25

polarographic wave should be identical in form with equation 57. Hence, just as in the electrodeposition of metal ions, the half-wave potential in the electroreduction of simple metal ions should correspond to the true inflection point of the wave and should be independent of the concentration of the reducible ions (44). In agreement with this prediction, the current-voltage curves obtained by Prajzler (95) show that the half-wave potential in the reduction of chromic to chromous ions is independent of the concentration of chromic ions.

2. Electrodeposition of complex metal ions

Although the deposition of metal ions from complex salts is much used in practical polarographic work (38, 50, 93, 106), the detailed kinetics of such reductions have not yet been thoroughly studied. Experimentally, it is found that the half-wave potentials of metal ions are shifted to more negative values by complex formation (figure 25). The half-wave potentials of complex metal ions are also generally independent of the concentration of the metal ion complex, provided an excess of the complex-forming substance is present (44).

Heyrovský and Ilkovič (44) assume that the reduction of complex metal ions takes place in two steps according to the scheme



They further assume that the dissociation of the complex proceeds rapidly enough to maintain equilibrium between the complex ions, the free metal ions, and the complex-forming anions at the surface of the dropping

FIG. 25. "Polarographic spectrum" according to V. Majer. Values of the half-wave potentials of inorganic ions in various media. I. Reduction and deposition potentials of cations in neutral or acidic solution. II. Reduction and deposition potentials of ions in alkaline solution. III. Reduction potentials of anions and molecules in neutral (solid line) or in acidic (dotted line) solution and in ammonia buffer (dashed line). NO_3^- and NO_2^- in 0.1 N lanthanum chloride. IV. Reduction potentials of complexes. Solid line, in 10 per cent sodium potassium tartrate; dashed line, in 1 N potassium cyanide. V. Depolarizing potentials of anions (anodic polarization). VI. Deposition potentials of cations of the commonly used indifferent electrolytes present in excess (1000 X) in the solution. VII. Potentials of the large mercury reference electrode in solutions containing the usual anions in ca. 1 N concentration. (The numbers 0, 1, and 2 refer to chloride concentrations of 1, 0.1, and 0.01 N.) All values are the potentials at the half diffusion current except those which are marked by a little circle signifying the contact point of a 45° tangent at a 2-cm. wave. The values of the potentials refer to the 1 N calomel electrode as zero and to room temperature.

electrode. If this assumption is correct, the concentration of free metal ions at the electrode surface should be given (44) by

$$(M^{n+}) = K_c(MX_b^{(b-n)-})/(X^{n-})^b \quad (63)$$

where K_c is the equilibrium constant (complex constant) of equation 61. On the basis of these assumptions Heyrovský and Ilkovič predict that the negative shift in the half-wave potential should be equal to $(RT/nF) \ln K_c$, when the concentration of the complex-forming anions, X^{n-} , is unity. It should be realized that this can be true only if the diffusion coefficient of the complex metal ions is the same as that of the simple (hydrated) metal ions. If the diffusion coefficients of the simple and complex metal ions are not equal, the negative shift in the half-wave potential will also depend on the logarithm of the ratio of the two diffusion coefficients, in addition to the logarithm of the complex constant. As yet these postulates have not been thoroughly tested experimentally.

The kinetics of the reduction of complex metal ions is undoubtedly much more complicated than the scheme represented by equations 61 and 62. In certain cases the rate of dissociation of the complex may be slower than the discharge of the liberated metal ions, and equilibrium will then not be maintained during the discharge. For example, Pines (93) has obtained evidence that the rate of dissociation of the various complex zinc cyanides is relatively slow, and that equilibrium is not maintained at the surface of the dropping electrode. Brochet and Petit (9) and Foerster (13), from studies of the alternating current electrolysis of complex metal cyanides, have obtained data which lead to the same conclusion. Furthermore, Herman (29) has obtained evidence that the dissociation of gold cyanide and gold hydroxide complexes and the rate of conversion of one complex into the other are relatively slow processes. If these interpretations are correct, it is evident that the actual concentration of free metal ions at the surface of the dropping electrode will be smaller than we would calculate from equation 54, and the rate of dissociation of the complex, rather than its rate of diffusion, should be the current-determining factor.

It is also conceivable that the reduction of very stable complex metal ions may take place by the direct capture of electrons from the electrode without preliminary dissociation. In such a case the rate of the direct interaction of the complex metal ions with electrons, and not the rate of dissociation of the complex, should be the current-determining factor on the polarographic wave. If the positively charged central ion in the complex is well shielded the capture of electrons may be hindered, and the potential of the dropping electrode will then not correspond to equilibrium conditions.

Tomeš (119) has studied the kinetics of the reduction of the weak electrolyte mercuric cyanide, a case quite similar to the reduction of complex metal ions. In this case the reduction apparently takes place reversibly. Tomeš found that the current-voltage curve was not symmetrical with respect to the half-wave point, and that consequently the half-wave potential was shifted to more negative values with increasing concentration of the mercuric cyanide (no excess cyanide present). He also found that the half-wave potential was greatly dependent on the pH of the solution, owing to the hydrolysis of the liberated cyanide ions at the surface of the mercury drops. Tomeš showed that these results correspond with the theory on the basis that the reduction takes place reversibly.

The discharge of hydrogen. The discharge of hydrogen ions at the dropping electrode has been the subject of numerous investigations, which are discussed by Professor Heyrovský in an accompanying paper in this Journal. Therefore we shall mention only a few of the more important results obtained. On the basis of theoretical consideration Heyrovský (32) predicted that the discharge potential (not $E_{1/2}$) of hydrogen ions should be shifted to more positive values with increasing hydrogen-ion concentration, according to

$$(2RT/F) \Delta \ln C_{H^+}$$

These predictions were experimentally verified by Herasymenko (25). Herasymenko and Šlendyk (27) have also shown that the overvoltage of hydrogen is shifted to more negative values with increasing concentration of neutral salts in the solution. They showed that the negative shift in the overvoltage was greater, the greater the valence of the cations of the neutral salt, but that in all cases the discharge potential approached a limiting constant value with increasing concentration of the neutral salt.

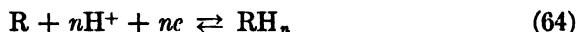
Tomeš (119) has recently made a careful study of the discharge of hydrogen from strong and weak acids. In the discharge of hydrogen from hydrochloric acid in the presence of excess lithium chloride or calcium chloride, he found that $E_{1/2}$ was not constant, but was shifted by 0.028 volt to more positive values for a tenfold increase in the concentration of hydrochloric acid. In the discharge of hydrogen ions from acetic acid in the presence of excess lithium chloride, Tomeš found that $E_{1/2}$ was shifted by 0.087 volt to more negative values per tenfold increase in the acetic acid concentration. He also found, in the discharge of hydrogen from buffer mixtures of acetic acid and lithium acetate in excess lithium chloride, that $E_{1/2}$ was shifted to more negative values with increasing concentration of the lithium acetate at a constant concentration of acetic acid, but that with a constant concentration of lithium acetate $E_{1/2}$ was shifted by 0.027

volt to more positive values per tenfold increase in the acetic acid concentration.

Heyrovský and Müller (45), and especially Novák (86), have studied the discharge of hydrogen and deuterium at the dropping electrode from mixtures of light and heavy water. Novák found that the half-wave potential of hydrogen from hydrochloric acid was 0.087 volt more negative in heavy water (99.6 per cent) than in ordinary water. The hydrogen overvoltage in light and heavy water and the separation coefficient of hydrogen and deuterium at the dropping mercury electrode have been exhaustively discussed by Heyrovský (39).

3. Reduction of organic compounds

According to Heyrovský and Ilkovič (44), the reversible reduction of organic compounds at the dropping electrode may be represented by the equation



In the majority of cases n is equal to 2. If this reduction takes place reversibly the potential of the dropping electrode should be given by

$$E = E^\circ - (RT/nF) \ln (C_{RH_n}^0/C_R^0) + (RT/F) \ln C_{H^+}^0 \quad (65)$$

If the solution originally contains none of the reduced form RH_n , we may assume that $C_{RH_n}^0 = ki$. Furthermore, C_R^0 at the surface of the dropping electrode should be related to the concentration of R in the body of the solution, C_R , by an expression analogous to equation 53, that is,

$$C_R^0 = C_R - i/K = (i_d - i)/K \quad (66)$$

where, as before, K is the diffusion current constant. Hence equation 65 becomes

$$E = E^\circ - (RT/nF) \ln (i/(i_d - i)) - (RT/nF) \ln kK + (RT/F) \ln C_{H^+}^0 \quad (67)$$

It is evident then that, if the reduction takes place reversibly, the current-voltage curve should be symmetrical about the half-wave point. Hence at a constant value of the hydrogen-ion concentration the half-wave potential should be constant and independent of the concentration of the reducible organic substance (44). This is found to be true in many cases (11b).

Furthermore, E° is the ordinary reversible oxidation reduction potential of the system. From equation 67 we see that the half-wave potential when $i = i_d/2$ should be given by

$$E_{1/2} = E^\circ - (RT/nF) \ln kK + (RT/F) \ln C_{H^+}^0 \quad (68)$$

According to Heyrovský and Ilkovič the product kK in this equation is equal to $(D_{\text{ox}}./D_{\text{red}}.)^{1/2}$, where $D_{\text{ox.}}$ and $D_{\text{red.}}$ are, respectively, the diffusion coefficients of the oxidized and reduced forms of the reducible substance. Since $(D_{\text{ox}}./D_{\text{red}}.)^{1/2}$ will usually have a value close to unity, the half-wave potential in reversible reductions should be practically coincident with the reversible oxidation-reduction potential of the system at a given pH. Müller and Baumberger (81) and Tachi (115) have demonstrated that this is true in the reduction of quinone, in the oxidation of hydroquinone, and in the reversible oxidation and reduction of various other organic substances at the dropping electrode.

The reduction of many organic substances at the dropping electrode does not take place reversibly; in fact, irreversible reductions are more common than reversible reductions. For example, the reduction of nitro compounds, of most unsaturated compounds, and of aldehydes and ketones is an irreversible process at the dropping electrode. According to Heyrovský and Ilkovič (44) and Tachi (115), even in such cases the half-wave potential is a characteristic constant, at a given pH, of the particular substance and is independent of its concentration. This has been demonstrated by Vopička (125) in the irreversible reduction of fumaric and maleic acids. In the irreversible reduction of these substances the negative shift in $E_{1/2}$ with increasing pH is usually different from what we would expect from equation 68. For example, Vopička found that the shift in the half-wave potential of maleic acid with increasing pH was twice as great, and that of fumaric acid three times as great, as would be expected from equation 68. This difference makes possible the differentiation of these two stereoisomers in alkaline medium.

The effect of the structure of organic compounds, particularly aldehydes and ketones, on their reduction potentials at the dropping electrode has been studied by Shikata and Tachi (108), Winkel and Proske (126), Adkins and Cox (3), and Tachi (115). The reduction potentials of the various isomeric dinitrophenols and dinitrobenzenes have been studied by Shikata and Hozaki (107). Tables of the reduction potentials of a variety of organic substances have also been given by Heyrovský (38) and Hohn (50). The reduction of organic compounds at the dropping electrode is discussed in detail by Dr. O. H. Müller in an accompanying paper in this Journal.

X. ANODIC CURRENT-VOLTAGE CURVES WITH THE DROPPING ELECTRODE

So far we have discussed the current-voltage curves obtained with the dropping electrode functioning as cathode, that is, electroreductions. It is also possible to obtain current-voltage curves, and limiting currents, with the dropping electrode functioning as the anode of the cell.

Revenda (99) has shown that limiting currents are obtained with the dropping mercury anode in dilute solutions of halide, cyanide, sulfide, thiocyanate, hydroxide, and other ions which form insoluble or complex compounds with mercury. In such cases the primary electrode reaction is the anodic dissolution of mercury,



In solutions containing no ions which form insoluble or complex mercury salts, this reaction sets in at the dropping anode at about +0.3 volt (99), and with increasing positive potential the corresponding current increases rapidly and apparently without limit. If the anodic dissolution of mercury takes place reversibly, the potential of the dropping anode should be given approximately by

$$E = 0.518 + (RT/2F) \ln C_{\text{Hg}_2^{++}}^0 \quad (70)$$

where 0.518 volt is the standard potential of the mercury electrode against the normal calomel electrode, and $C_{\text{Hg}_2^{++}}^0$ is the concentration of mercurous ions formed at the surface of the mercury drops. If the solution contains ions which form insoluble or complex mercury salts, the dropping anode will be depolarized and the current will increase rapidly when the potential is made sufficiently positive so that the concentration of mercurous ions at the surface exceeds the solubility product, or the complex constant, of the particular insoluble or complex mercury salt. For example, if the solution contains chloride ions, the dropping anode will be depolarized when the concentration of mercurous ions at the electrode surface exceeds the value $S/(C_{\text{Cl}}^0)^2$, where S is the solubility product of mercurous chloride (6×10^{-19}) and C_{Cl}^0 is the concentration of chloride ions at the electrode surface. Hence from equation 70 we calculate that in a 0.001 N chloride solution the dropping electrode should become depolarized and the current should increase at a potential of about +0.16 volt. This is in good agreement with the experimental value of +0.17 volt found by Revenda.

When the dropping electrode is depolarized by chloride ions, the observed current is a measure of the rate of precipitation of mercurous chloride at the electrode surface. This rate, in turn, depends on the rate of diffusion of chloride ions up to the electrode from the body of the solution. As the potential is made more and more positive, the rate of diffusion of chloride ions approaches a constant value, and hence a limiting current results. The same principles apply in the case of other ions which form insoluble or complex salts with mercury. Revenda found that these anodic limiting currents were proportional to the concentration of the depolarizing ion, and that when several depolarizing ions were present in the solution each one produced its own characteristic wave on the current-voltage curve.

The potential at which a given ion at a given concentration depolarizes the dropping anode is, of course, more negative the more insoluble, or the more complex, the resulting mercurous salt is. The depolarization potentials of various ions, determined by Revenda, are indicated in figure 25.

It is also possible for the dropping electrode to function simply as an indifferent anode in certain (chiefly organic) oxidation reactions. For example, Müller and Baumberger (81) have shown that hydroquinone can be oxidized to quinone to give a limiting current at the dropping anode. The range of potentials over which the dropping electrode can function as an indifferent anode is limited, of course, by the potential at which the dropping electrode becomes depolarized by depolarizing ions that may be present in the solution. For example, in a pure solution of potassium nitrate the dropping electrode can function as an indifferent anode up to a potential of about +0.3 volt, in solutions containing chloride ions up to a potential of about +0.2 volt, but in solutions containing sulfide ions only at potentials more negative than about -0.7 volt. In this respect the dropping mercury anode is quite limited in its application and less suitable than a platinum microanode. It should be possible to use a platinum microanode up to the potential at which oxygen evolution begins, that is, up to about +1.0 volt.

XI. ANALYTICAL APPLICATIONS OF THE POLAROGRAPHIC METHOD AND POLAROMETRIC TITRATIONS

1. General applications

It is impossible in the space available to give a detailed discussion of the many applications of the polarographic method in the various branches of chemistry and allied fields. We shall, therefore, confine the present brief review to a mere outline of practical polarographic analysis and to a few examples which demonstrate the versatility of the method. For more extensive discussions of practical details the reader is referred to the monographs of Heyrovský (38), Semerano (104), and particularly that of Hohn (50), and the papers of Hohn (51), Kemula (60), Thanheiser and Maassen (117), Maassen (69), and Winkel and Proske (127). The applications of the method in microanalysis have been discussed by Heyrovský (34), Hamamoto (21), and Heller, Kuhla, and Machek (24). Details of special apparatus for micropolarographic analysis, with which a volume of solution as small as 0.005 cc. may be analyzed, have been discussed by Majer (74). Applications of polarographic analysis in physiological chemistry and medicine have been made by Tropp (120), Scuberting (105), Rosenthal (100), Mandai (77), Petráček (91), and Brdička (6).

We have seen that, under optimum conditions, the half-wave potential is a characteristic constant for each electroreducible substance (qualitative

analysis), while the diffusion current is proportional to its concentration (quantitative analysis). In general, the optimum conditions for practical polarographic analysis are as follows: (a) The migration current of reducible ions should be eliminated by having an excess of an indifferent salt present in the solution. (b) Maxima should be eliminated. (c) The diffusion current should be well defined, and preferably directly proportional to the concentration of the reducible substance (after proper correction for the residual current). (d) The concentration of foreign or indifferent electrolytes should be large enough so that a small variation in their concentration will not appreciably affect the diffusion current. (e) Adsorption of electroreducible ions (particularly heavy-metal ions) by the glass wall of the cell must be prevented. Such adsorption may be particularly marked when dealing with very small concentrations of heavy-metal ions (e.g., micropolarographic determination of lead ions), but it can usually be prevented by acidifying the solution. (f) Since diffusion current constants vary with the temperature, the latter should be controlled to at least $\pm 0.5^{\circ}\text{C}$.

In general the most suitable medium (i.e., nature and concentration of foreign electrolytes) for the polarographic determination of a given substance must be determined by trial. In order to eliminate maxima and obtain well-defined diffusion currents and to insure a practically constant composition of the solution with respect to foreign electrolytes, Hohn (50, 51) recommends the use of so-called "Grundlösungen" or "regulating solutions." These are composed of suitable salts, or salt mixtures, with the addition, when necessary, of substances which will suppress maxima. Some of the regulating solutions contain complex-forming salts, buffer mixtures, or reagents which will precipitate, or otherwise remove, interfering ions from the solution to be analyzed. The solution to be analyzed is added to about five or ten times its volume of the proper regulating solution. The most favorable composition of the regulating solution for each substance, or group of substances, to be determined must be established by preliminary experiments. Hohn (50) has given a list of regulating solutions that he found to be most suitable for the determination of various metal ions. The reader is referred to his monograph and also to the papers of Maassen (69) and Thanheiser and Maassen (117) for details of their use.

The residual current of each regulating solution should be determined in a "blank" experiment, and the value obtained subtracted from the observed diffusion currents in this particular regulating solution to obtain the true diffusion currents.

When the (corrected) diffusion current of a given substance is directly proportional to its concentration, the diffusion current constant of the

substance in the particular regulating solution used may be determined in preliminary experiments with known amounts of the substance. The concentration of this substance in the solution analyzed is thus obtained simply by dividing the (corrected) diffusion current of the substance by the diffusion current constant. When the diffusion current is not directly proportional to the concentration, an empirical calibration curve may be used.

The concentration of the particular substance responsible for a given wave on an "unknown" polarogram can also be determined in the following manner (50, 51). The polarogram of the "unknown" solution is first obtained in the usual way. A known amount of the substance in question is then added to the solution, and a second polarogram is obtained. From the increase (corrected for the residual current) of the diffusion current caused by the known addition, the original concentration of the substance can be easily computed by simple proportion. The particular advantage of this method is that the "calibration" (i.e., the diffusion current constant) is obtained under identically the same conditions with which the original polarogram was obtained, and accidental differences in composition are minimized. Obviously, this "internal standard" method can be used only when the diffusion current is directly proportional to the concentration.

When dealing with mixtures of electroreducible substances, it is often possible to detect and determine several of the constituents from a single current-voltage curve (figure 4). There is no difficulty in detecting and determining a trace of one substance in the presence of a large excess of another electroreducible substance when the former is reduced at a more positive potential than the latter. On the other hand, when a large excess of a more easily reducible substance is present in the solution of the substance to be determined, the *direct* determination of the microconstituent becomes very difficult and often almost impossible. For example, it is easy to detect directly and to determine a trace of lead or cadmium in zinc or zinc salts, since the former ions are reduced at a more positive potential than zinc ions (figure 25), but it is impossible to determine directly a trace of zinc or cadmium in lead salts without previous separation. In cases like the latter, Ilković and Semerano (57) recommend that the diffusion current of the interfering major constituent be "compensated" by sending a current of equal magnitude through the current-measuring galvanometer in an opposite direction from an outside source. In this way the wave of the microconstituent can be obtained on the polarogram, but the current oscillations in this case are usually so large that the diffusion current cannot be determined accurately.

In order to determine a microconstituent in the presence of a large excess of a more easily reducible substance, it is usually necessary to resort

to preliminary chemical or physical methods of separation. For instance, a trace of zinc in lead salts can be determined by first precipitating the lead with sulfate and then determining the zinc in the filtrate. It is often not necessary to filter off the precipitate of the interfering major constituent. Furthermore, the precipitation of the major constituent need not be quantitatively complete; it is usually necessary only to reduce its concentration to a value of about the same order of magnitude as that of the microconstituent to be determined. When such precipitation methods are used to remove interfering ions the possibility of coprecipitation of the microconstituent must always be kept in mind.

When the half-wave potentials of two electroreducible metal ions are close together, their waves tend to overlap. Many such combinations may occur, as shown by the "polarographic spectrum" in figure 25. In such cases it is often possible to separate the two waves by adding a reagent, or by employing a regulating solution which contains a reagent, that forms a stable complex with one of the constituents but forms no complex, or a much less stable one, with the other. For example, the polarogram of a solution of potassium chloride containing lead and thallous ions shows only a single wave, but in the presence of potassium cyanide, tartrate, or an excess of hydroxide ions, the half-wave potential of lead is shifted to a more negative value while that of thallium remains virtually unchanged (figure 25). The result is the appearance of two separate waves, from which the concentrations of the lead and thallium can be found. If the single wave due to two substances cannot be resolved by complex formation, classical chemical methods of separation must be employed. It will be obvious that in the case of overlapping waves the previous chemical separation must be quantitatively complete.

Overlapping waves of organic substances can sometimes be resolved, or separated, by suitable adjustment of the pH of the solution. For example, in acid medium the waves due to a mixture of fumaric and maleic acids nearly overlap, but in alkaline medium two separate waves are obtained.

It should also be mentioned that the diffusion current of a given substance may, in certain cases, be influenced by the preceding discharge of another substance. For example, Lingane (65) found that the diffusion current of iodate was greatly decreased by the preceding discharge of thallous or cadmium ions. This peculiar effect has not yet been explained and is being further investigated.

Regarding the precision and accuracy of polarographic analysis the following brief statement may be made. Diffusion currents are easily reproducible within 1 per cent. When the proper correction for the residual current is applied, the concentration of a reducible substance can be determined within about ± 2 per cent when its concentration is greater

than $10^{-4} M$ and within about ± 10 per cent when the concentration is between 10^{-4} and $10^{-5} M$. On the basis of our experience we do not recommend the polarographic method for concentrations smaller than about $10^{-5} M$.

A unique advantage of the polarographic method is that the composition of the solution remains virtually unchanged, because the quantity of electricity passed through it is so very small. Hence after the electrolysis the sample can be used for other purposes, if desired.

One of the most useful applications of the polarographic method is the detection and determination of reducible metal ions in various substances. Typical examples of such applications are the determination of traces of lead in citric acid and the determination of traces of copper, lead, and cadmium in commercial zinc and zinc salts (38). Hohn (50, 51) has developed procedures for the rapid determination of copper, zinc, iron, lead, and nickel in brass and similar alloys. Thanheiser and Maassen (117) and Maassen (69) have applied the polarographic method to the determination of copper, nickel, and cobalt in steel. A rapid method for the polarographic determination of aluminum, zinc, manganese, and lead in commercially important magnesium alloys has been developed by Gull (18). Knoke (62) employed the polarographic method for the determination of zinc oxide in lithopone. The possibility of polarographically differentiating various rare-earth metals in mixtures has been studied by Noddack and Bruckl (85). Leach and Terry (64) have employed the polarographic method for the determination of scandium in aluminum salts. A simple method for the rapid determination of sodium and potassium, after preliminary separation by the perchlorate method, has been developed by Abresch (2). The polarographic determination of the alkali metals has also been studied by Majer (73), Heyrovský and Bureš (42), Lohnis, Meloche, and Juday (67), and Perrachio and Meloche (89). The polarographic microdetermination of manganese in biological material has been studied by Hamamoto (21). The use of "dithizone" for group separations of metal ions prior to their polarographic determination in mineral waters has been studied by Heller, Kuhla, and Machek (24). Preliminary "dithizone" separations have also been used by Stout and his collaborators (114) for the polarographic determination of small amounts of iron, manganese, copper, bismuth, lead, cadmium, zinc, nickel, and cobalt in plant ash. These few examples illustrate the usefulness of the polarographic method for the determination of metal ions in a variety of different materials.

Various oxygen-containing anions are reducible at the dropping electrode and may be polarographically determined. As examples, we may mention nitrate and nitrite (118), molybdate (121), bromate and iodate (101),

and selenite and tellurite ions (103). It should be mentioned, however, that these reductions are irreversible, and that the shapes of the current-voltage curves are greatly influenced by the composition of the electrolysis medium, especially by traces of polyvalent cations such as lanthanum ions (118). Hence in practical work the composition of the medium must be carefully controlled (by a regulating solution).

It has already been mentioned that oxygen is reducible at the dropping electrode. Vítek (124) has made use of this fact for the determination of oxygen in various technical gases. Petering and Daniels (90) have recently employed the polarographic determination of oxygen for studying the photosynthesis and respiration of green algae (*Chlorella pyrenoidosa*) and the respiration rate of yeast, blood cells, and animal tissue. According to these authors the polarographic method is superior to the classical manometric and Winkler methods for investigations of this kind, especially when the oxygen uptake is very small.

The suppression of the prominent oxygen maximum by various capillary-active substances, such as dyes (98), lyophilic colloids (122), fatty acids (97), and alkaloids (20), may be used for determining traces of such substances. For example, Gosman and Heyrovský (17) determined certain capillary-active impurities in petroleum and its distillates by their suppressive effect on the oxygen maximum. Heyrovský, Smoleř, and Šťastný (38) employed the same principle for differentiating between natural and synthetic vinegar.

It has been mentioned that substances which decrease the overvoltage of hydrogen on mercury catalyze the evolution of hydrogen and cause the appearance of a catalytic wave at potentials more positive than the "normal" discharge potential of hydrogen. For example, the discharge of hydrogen from strong acids is catalyzed by traces of noble-metal ions, such as platinum, ruthenium, and palladium (112, 111). This phenomenon may be utilized for the detection and determination of very minute amounts of these noble-metal ions. The distinguishing characteristic of such catalytic waves is that they are ten to twenty times larger than the waves due to simple reduction of the same concentration of the catalyst ion. Perhenate ions in acetate buffers containing hydrogen sulfide also catalyze the discharge of hydrogen and cause the appearance of a catalytic wave. Heyrovský (37) utilized this fact for detecting traces of rhenium in maunganous salts. The discharge of hydrogen is also catalyzed by various quinoline derivatives, such as quinine, quinidine, cinchonine, and cinchonidine (88).

Various proteins and organic compounds containing a sulphydryl group (e.g., cystine, cysteine) catalyze the evolution of hydrogen from the ammonium ion in ammonia-ammonium chloride buffers containing a small

concentration of cobalt or nickel ions (6, 40). The presence of cobalt or nickel ions is essential for the production of a catalytic wave. Proteins give the catalytic wave with either cobaltous or cobaltic ions, but with sulfhydryl compounds a catalytic wave is obtained only with cobaltous and not with cobaltic ions (differentiation between proteins and catalytic sulfhydryl compounds). Brdička (6, 8) has utilized the catalytic effects due to the sulfhydryl groups in denatured blood serum for the clinical diagnosis of cancer. According to his results the catalytic effect is smaller with carcinomatic than with non-carcinomatic serum. Rosenthal (100) has described several other practical applications of these catalytic effects in biological chemistry.

The polarographic method gives promise of becoming a valuable tool in organic chemistry. A large variety of organic substances, including various aldehydes and ketones, nitro compounds, azo and diazo compounds, quinones, and simple unsaturates, are reduced and give well-defined waves with the dropping electrode. Applications of the polarographic method in organic chemistry are discussed in detail by Dr. O. H. Müller in an accompanying paper in this Journal.

From the foregoing brief review it will be evident that the polarographic method is rapidly attaining an important place in quantitative and qualitative analysis. However, much fundamental and systematic research still remains to be done before the method can be used reliably and accurately for general analytical purposes. More systematic investigations of the effect of the nature of the solvent on diffusion currents, the elimination of maxima, and irreversible oxidations and reductions are particularly needed. In order to avoid possible disappointments, we wish to warn practical analytical chemists that the method is still in the developmental stage. When dealing with complex and unfamiliar mixtures, it is imperative to test the reliability of the method with synthetic mixtures of composition closely comparable to that of the unknown.

2. Polarometric titrations

A new method of electrometric titration employing the dropping electrode has recently been introduced by Heyrovský (33) and Heyrovský and Berezický (41). The method is based on the decrease of the diffusion current of a reducible substance when it is removed from solution by titration with a suitable reagent or on the sudden increase in current at the equivalence point when the reagent is titrated with a solution of a reducible substance. The general principles of the method have recently been discussed by Majer (75), who coined the name "polarometric titration" for it and studied the direct and reverse titration of lead with sulfate.

The application of polarization phenomena to titrations is not new.

The "dead-stop end point" method of Foulk and Bawden (14) and the method of Guzmán and Rancaño (19) are based on the sudden polarization or depolarization of an indicator electrode at the equivalence point and a corresponding abrupt change in current. However, these methods are based on "chemical polarization," whereas polarometric titrations with the dropping electrode are based on concentration polarization. Hardly any work has been done as yet to introduce polarometric titrations into general volumetric practice. At the present time, systematic studies of the method are being conducted in this laboratory, and from the results obtained so far we predict that in the future polarometric titrations will become as important and useful in volumetric analysis as potentiometric titration methods.

Since we are planning to publish more extensive papers on the subject in the near future and space in this paper is limited, we shall confine the present discussion to the basic principles of polarometric titrations, and a few examples which indicate its possibilities. Suppose that we titrate a dilute solution of lead ions with oxalate in a polarographic cell, with the applied E.M.F. maintained constant at a value such that the diffusion current of lead is obtained, and measure the diffusion current after each addition of the standard oxalate solution. The titration curve will be as shown schematically in figure 26a. The diffusion current decreases continuously during the titration as the lead ions are removed by precipitation as lead oxalate. When the equivalence point is reached, and the precipitation of lead is fairly complete, the current becomes constant and independent of further addition of oxalate. When the solubility of the precipitate is negligibly small the break at the equivalence point will be sharp, and the final constant current will be very small and practically equal to the condenser current. If the solubility of the precipitate at the equivalence point is not negligible, the titration curve will be more or less curved in the region of the equivalence point. In such a case (e.g., titration of Pb^{++} with SO_4^{-}) the end point is found from the intersection of the extrapolated straight portions of the curve after correction for the dilution effect, much in the same way as the end points of conductometric titrations are determined. If the solution being titrated contains another reducible substance which is reduced at the particular value of the applied E.M.F. but which is not precipitated or otherwise removed by the reagent employed, the final constant current will be almost equal to the diffusion current of this substance. For example, this would be the case if a solution of lead were titrated with oxalate without removing air from the solution.

Suppose that the ion to be determined is not electroreducible at the particular value of the applied E.M.F. used, but the reagent ion is. This is the case, for example, when barium is titrated with chromate in neutral

solution with the applied E.M.F. at 1.5 volts or less. The curve of this titration is shown schematically in figure 26b.

The sharpest break at the equivalence point, and the most accurate results, are to be expected when the ion being titrated and the reagent ion are both reducible at the particular applied E.M.F. In such a case the titration curve will be V-shaped. For example, in the titration of lead ions with chromate the titration curve is as shown schematically in figure 27a, because both lead and chromate ions are reducible at the dropping electrode.

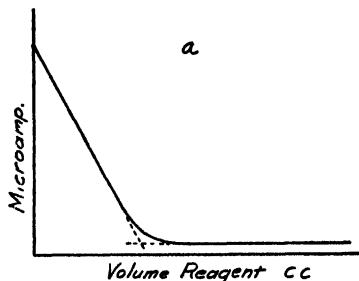


FIG. 26

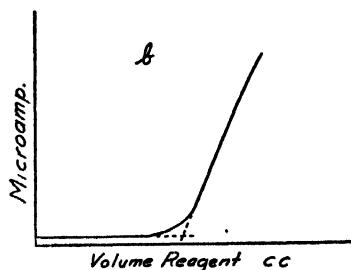


FIG. 26

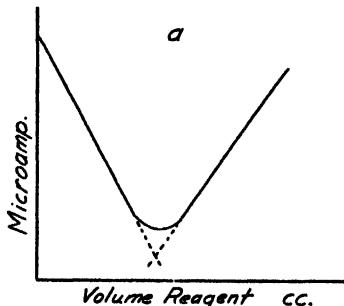


FIG. 27

FIG. 26. Schematic representation of polarometric titration curves
FIG. 27. Schematic representation of polarometric titration curves

The foregoing titrations have actually been performed in this laboratory with excellent results. The polarometric titration method promises to be particularly valuable in cases for which reversible indicator electrodes are not available for titration by the classical potentiometric method. Polarometric titrations are also especially suitable for very small concentrations. For example, a 10^{-4} molar lead solution can easily be titrated with chromate.

It is also to be expected that use can be made of fractional precipitation for determining two substances by a single titration. For example, if the

solubilities of lead chromate and barium chromate are sufficiently different, and if barium ions do not coprecipitate with lead chromate, it should be possible to titrate a solution containing lead and barium ions with chromate. If the foregoing conditions are fulfilled, the titration curve should be as shown schematically in figure 27b. The first break corresponds to the complete precipitation of lead and the second to that of barium.

Many other types of polarometric titrations are possible. We have titrated chloride ions with silver by employing the dropping electrode as anode (see section X). In this case the current is originally negative, decreases to zero at the equivalence point, and then becomes positive when an excess of silver ions has been added. The titration curve is practically a straight line, and the equivalence point is indicated by zero current. Under suitable conditions various oxidation-reduction and complex-formation reactions can also be made the basis of polarometric titration methods. In future publications limitations of the polarometric titration method will be discussed.

XII. SUMMARY

1. The fundamental principles of polarography developed by Heyrovský and his school have been discussed in detail. Applications of the method to problems in physical, organic, and analytical chemistry, and in allied fields have been pointed out.

- 2 From the analytical viewpoint much systematic work remains to be done before the polarographic method can be generally applied. It is anticipated that polarometric titrations will become an important method in volumetric analysis

3. The "apparent diffusion current" is the sum of the true diffusion current and the condenser current.

4. The factors determining the occurrence of maxima are involved and not completely understood.

5. Quite generally an electroreducible substance is characterized by its constant "half-wave potential." When the electroreduction on the dropping mercury electrode is reversible, a mathematical analysis of the polarographic wave yields the number of electrons involved in the electro-reduction.

XIII. REFERENCES⁴

- (1) ABEGG, R., AND BOSE, E : Z physik. Chem. **30**, 545 (1899).
- (2) ARRENCHE, K : Angew. Chem. **48**, 683 (1935); Chem. Fabrik **8**, 380 (1935)
- (3) ADKINS, H., AND COX, FRED W : J Am. Chem. Soc. **60**, 1151 (1938).

⁴ A complete bibliography (including titles of the papers) of the literature on the polarograph up through 1937 has been published by J. Heyrovský and J. Klumper (Collection Czechoslov. Chem. Commun. **10**, 153 (1938)).

- (4) ANTWEILER, H. J.: *Z. Elektrochem.* **43**, 596 (1937); **44**, 719 (1938).
- (5) BORCHERDT, G. T., MELOCHE, V. W., AND ADKINS, H.: *J. Am. Chem. Soc.* **59**, 2171 (1937).
- (6) BRDIČKA, R.: Collection *Czechoslov. Chem. Commun.* **5**, 112, 148, 238 (1933); *Mikrochemie* **15**, 167 (1934); Collection *Czechoslov. Chem. Commun.* **8**, 366 (1936); *Nature* **139**, 330, 1020 (1937).
- (7) BRDIČKA, R.: Collection *Czechoslov. Chem. Commun.* **8**, 419 (1936).
- (8) BRDIČKA, R.: *Acta intern. Verein. Krebsverkämpfung* **3**, 13 (1938); *Wochschr. prakt. "Arzte"* *Medizinische Klinik* **33**, 1186 (1937).
- (9) BROCHET, A., AND PETIT, F.: *Z. Elektrochem.* **10**, 909 (1904).
- (10) BRUNS, B., AND FRUMKIN, A.: *Acta Physicochim. U. R. S. S.* **1**, 232 (1934).
- (11) EMELIANOVA, N. V., AND HEYROVSKÝ, J.: *Trans. Faraday Soc.* **24**, 257 (1928).
- (12) ERDEY-GRÚZ, T., AND VOLMER, M.: *Z. physik. Chem.* **A157**, 165 (1931).
- (13) FOERSTER, F.: *Z. Elektrochem.* **13**, 561 (1907).
- (14) FOULK, C. W., AND BAWDEN, A. T.: *J. Am. Chem. Soc.* **48**, 2045 (1926).
- (15) FRUMKIN, A.: *Ergeb. exakt. Naturw.* **7**, 235-75 (1928). Die Elektrokapillarkurve.
- (16) FRUMKIN, A., AND BRUNS, B.: *Acta Physicochim. U. R. S. S.* **1**, 232 (1934).
- (17) GOSMAN, B., AND HEYROVSKÝ, J.: *Trans. Electrochem. Soc.* **59**, 249 (1931).
- (18) GULL, H. C.: *J. Soc. Chem. Ind.* **56**, 177 (1937).
- (19) GUZMÁN, J., AND RANCAÑO, A.: *Anales soc. españ. físi. quím.* **32**, 590 (1934).
- (20) HAMAMOTO, E.: Collection *Czechoslov. Chem. Commun.* **5**, 427 (1933).
- (21) HAMAMOTO, E.: Collection *Czechoslov. Chem. Commun.* **6**, 325 (1934).
- (22) HARTLEY, G. S.: *Phil. Mag.* **12**, 473 (1931).
- (23) HASKELL, R.: *Phys. Rev.* **27**, 145 (1908).
- (24) HELLER, K., KUHLA, G., AND MACHEK, F.: *Mikrochemie* **18**, 193 (1935); **23**, 78 (1937).
- (25) HERASYMENKO, P.: *Rec. trav. chim.* **44**, 503 (1925).
- (26) HERASYMENKO, P.: *Trans. Faraday Soc.* **24**, 257 (1928).
- (27) HERASYMENKO, P., AND ŠLENDÝK, I.: *Z. physik. Chem.* **A149**, 123 (1930).
- (28) HERASYMENKO, P., HEYROVSKÝ, J., AND TANČAKIJKÝ, K.: *Trans. Faraday Soc.* **25**, 152 (1929).
- (29) HERMAN, J.: Collection *Czechoslov. Chem. Commun.* **6**, 37 (1934).
- (30) HERMANS, J. J.: Diffusion of Electrolytes, Dissertation, Leiden, 1937; *Rec. trav. chim.* **56**, 635 (1937).
- (31) HEYROVSKÝ, J.: *Chem. Listy* **16**, 256-304 (1922); *Phil. Mag.* **45**, 303-14 (1923).
- (32) HEYROVSKÝ, J.: *Rec. trav. chim.* **44**, 499 (1925).
- (33) HEYROVSKÝ, J.: *Bull. soc. chim.* **41**, 1224 (1927).
- (34) HEYROVSKÝ, J.: *Mikrochemie* **12**, 25 (1932).
- (35) HEYROVSKÝ, J.: A Polarographic Study of the Electrokinetic Phenomena of Adsorption, Electroreduction, and Overpotential Displayed at the Dripping Mercury Cathode; Actualités scientifiques et industrielles, No. 90. Hermann et Cie., Paris (1934).
- (36) HEYROVSKÝ, J.: *Archiv Hem. Farm.* **8**, 11 (1934).
- (37) HEYROVSKÝ, J.: *Nature* **135**, 870 (1935).
- (38) HEYROVSKÝ, J.: Polarographie, in W. Böttger's *Die physikalischen Methoden der chemischen Analyse*, Vol. 2, pp. 260-322. Akademische Verlagsgesellschaft, Leipzig (1936).
- (39) HEYROVSKÝ, J.: Collection *Czechoslov. Chem. Commun.* **9**, 273, 345 (1937).
- (40) HEYROVSKÝ, J., AND BABIČKA, J.: *Chem. News* **141**, 369, 385 (1930); Collection *Czechoslov. Chem. Commun.* **2**, 370 (1930).

- (41) HEYROVSKÝ, J., AND BEREZICKÝ, S.: Collection Czechoslov. Chem. Commun. **1**, 19 (1929).
- (42) HEYROVSKÝ, J., AND BUREŠ, M.: Collection Czechoslov. Chem. Commun. **8**, 446 (1936).
- (43) HEYROVSKÝ, J., AND DILLINGER, M.: Collection Czechoslov. Chem. Commun. **2**, 626 (1930).
- (44) HEYROVSKÝ, J., AND ILKOVIČ, D.: Collection Czechoslov. Chem. Commun. **7**, 198 (1935).
- (45) HEYROVSKÝ, J., AND MÜLLER, O. H.: Collection Czechoslov. Chem. Commun. **7**, 281 (1935).
- (46) HEYROVSKÝ, J., AND SHIKATA, M.: Rec. trav. chim. **44**, 496 (1925).
- (47) HEYROVSKÝ, J., AND SIMUNEK, R.: Phil. Mag. **7**, 951 (1929).
- (48) HEYROVSKÝ, J., AND VASCAUTZANU, E.: Collection Czechoslov. Chem. Commun. **3**, 418 (1931).
- (49) HOEKSTRA, J.: Collection Czechoslov. Chem. Commun. **6**, 17 (1934).
- (50) HOHN, H.: Chemische Analysen mit dem Polarographen. Julius Springer, Berlin (1937).
- (51) HOHN, H.: Z. Elektrochem. **43**, 127 (1937).
- (52) ILKOVIČ, D.: Collection Czechoslov. Chem. Commun. **4**, 480 (1932).
- (53) ILKOVIČ, D.: Collection Czechoslov. Chem. Commun. **6**, 498 (1934).
- (54) ILKOVIČ, D.: Collection Czechoslov. Chem. Commun. **8**, 13 (1936).
- (55) ILKOVIČ, D.: Collection Czechoslov. Chem. Commun. **8**, 170 (1936).
- (56) ILKOVIČ, D.: Collection Czechoslov. Chem. Commun. **10**, 249 (1938).
- (57) ILKOVIČ, D., AND SEMERANO, G.: Collection Czechoslov. Chem. Commun. **4**, 176 (1932).
- (58) IRVING, G. W., JR., AND SMITH, N. R.: Ind. Eng. Chem., Anal. Ed. **6**, 480 (1934).
- (59) JANDER, G., AND WINKEL, A.: Z. physik. Chem. A**149**, 97 (1930).
- (60) KEMULA, W.: Z. Elektrochem. **37**, 779 (1931).
- (61) KEMULA, W.: Congr. intern. quím. pura aplicada, 9th Congr., Madrid, 1934.
- (62) KNOE, S.: Angew. Chem. **50**, 728 (1937).
- (63) KUCERA, G.: Ann. Physik **11**, 529 (1903).
- (64) LEACH, R. H., AND TERRY, H.: Trans. Faraday Soc. **33**, 480 (1937).
- (65) LINGANE, J. J.: A Study of the Fundamental Principles of Electrolysis with the Dropping Mercury Electrode, Ph.D. Thesis, University of Minnesota, 1938.
- (66) LINGANE, J. J., AND KOLTHOFF, I. M.: J. Am. Chem. Soc. **61** (1939), in press.
- (67) LOHNIS, D., MELOCHE, V. W., AND JUDAY, C.: Trans. Wisconsin Acad. Sci. **31**, 285 (1938); Chem. Abstracts **32**, 9350 (1938).
- (68) MAAS, J.: De Polarografische Methode met de druppelende kwikelectrode ten dienste van het Pharmaceutisch Onderzoek, Dissertation, Amsterdam, 1937; Collection Czechoslov. Chem. Commun. **10**, 42 (1938).
- (69) MAASSEN, G.: Angew. Chem. **50**, 375 (1937); Physikalische Methoden in chemischen Laboratorium, Verlag Chemie, Berlin (1937).
- (70) MACDOUGALL, F. H.: Physical Chemistry. The MacMillan Co., New York (1936).
- (71) MACGILLAVRY, D.: Rec. trav. chim. **56**, 1039 (1937); **57**, 33 (1938).
- (72) MACGILLAVRY, D., AND RIDEAL, E. K.: Rec. trav. chim. **56**, 1013 (1937).
- (73) MAJER, V.: Z. anal. Chem. **92**, 321 (1933).
- (74) MAJER, V.: Mikrochemie **18**, 74 (1935).

- (75) MAJER, V.: *Z. Elektrochem.* **42**, 120, 123 (1936).
- (76) MAJER, V.: *Collection Czechoslov. Chem. Commun.* **7**, 146, 215 (1935); **9**, 360 (1937).
- (77) MANDI, H.: *Acta Schol. Med. Univ. Imp. Kioto* **14**, 163, 167 (1931).
- (78) MATHESON, L. A., AND NICHOLS, N.: *Trans. Electrochem. Soc.* **78**, 193 (1938).
- (79) MCBAIN, J. W., AND LIU, T. H.: *J. Am. Chem. Soc.* **53**, 59 (1931).
- (80) MÜLLER, O. H.: *Chem. Rev.* **24**, 95 (1939).
- (81) MÜLLER, O. H., AND BAUMBERGER, J. P.: *Trans. Electrochem. Soc.* **71**, 169, 181 (1937).
- (82) MÜLLER, R. H., GARMAN, R. L., DROZ, M. E., AND PETRAS, J.: *Ind. Eng. Chem., Anal. Ed.* **10**, 339 (1938).
- (83) NEJEDLÝ, V.: *Collection Czechoslov. Chem. Commun.* **1**, 319 (1929).
- (84) NERNST, W.: *Z. physik. Chem.* **2**, 613 (1888).
- (85) NODDACK, W., AND BRUCKL, A.: *Angew. Chem.* **49**, 533 (1936); **50**, 362 (1937).
- (86) NOVÁK, J.: *Collection Czechoslov. Chem. Commun.* **9**, 207 (1937).
- (87) ONSAGER, L., AND FUOSS, R. M.: *J. Phys. Chem.* **36**, 2689 (1932).
- (88) PECK, J.: *Collection Czechoslov. Chem. Commun.* **6**, 126, 190 (1934).
- (89) PERACCHIO, E. S., AND MELOCHE, V. W.: *J. Am. Chem. Soc.* **60**, 1770 (1938).
- (90) PETERING, H. G., AND DANIELS, F. J.: *Am. Chem. Soc.* **60**, 2796 (1938).
- (91) PETRÁČEK, E.: *Compt. rend., VIII congr. intern. dermatologie*, Copenhagen (1930), p. 813.
- (92) PHILPOT, J. ST. L.: *Phil. Mag.* [7] **13**, 775 (1932).
- (93) PINES, I.: *Collection Czechoslov. Chem. Commun.* **1**, 387, 429 (1929); *Chem. News* **139**, 196 (1929).
- (94) PLANCK, M.: *Wied. Ann.* **39**, 161, 561 (1890); *Sitzb. preuss. Akad. Wiss.* **285** (1927); **9** (1929); **367** (1930); **113** (1931).
- (95) PRAJZLER, J.: *Collection Czechoslov. Chem. Commun.* **3**, 406 (1931).
- (96) PROSKURNIN, M., AND FRUMKIN, A.: *Trans. Faraday Soc.* **31**, 110 (1935).
- (97) RASCH, J.: *Collection Czechoslov. Chem. Commun.* **1**, 560 (1929).
- (98) RAÝMAN, B.: *Collection Czechoslov. Chem. Commun.* **8**, 314 (1931).
- (99) REVENDA, J.: *Collection Czechoslov. Chem. Commun.* **6**, 453 (1934).
- (100) ROSENTHAL, H. G.: *Mikrochemie* **22**, 233 (1937).
- (101) RYLICH, A.: *Collection Czechoslov. Chem. Commun.* **7**, 288 (1935).
- (102) SCHMIDT, A.: *Z. Elektrochem.* **44**, 699 (1938).
- (103) SCHWAER, L., AND SUCHÝ, K.: *Collection Czechoslov. Chem. Commun.* **7**, 26 (1935).
- (104) SEMERANO, G.: *Il polarografo, sua teoria e applicazioni*, 2nd edition. Draghi, Padova (1933).
- (105) SEUBERLING, O.: *Klin. Wochschr.* **18**, 644 (1937).
- (106) SHIKATA, E., AND KIDA, Y.: *J. Phys. Chem. (Japan)* **2**, 75 (1929).
- (107) SHIKATA, M., AND HOZAKI, N.: *Mem. Coll. Agr. Kyoto Imp. Univ.*, No. **17**, 1, 21 (1931).
- (108) SHIKATA, M., AND TACHI, I.: *Mem. Coll. Agr. Kyoto Imp. Univ.*, No. **8**, 1-19 (1930); *Bull. Agr. Chem. Soc. Japan* **4**, 44 (1928).
- (109) SIEBERT, H., AND LANGE, T.: *Chem. Fabrik.* **11**, 141 (1938).
- (110) ŠLENDÝK, I.: *Collection Czechoslov. Chem. Commun.* **8**, 385 (1931).
- (111) ŠLENDÝK, I.: *Collection Czechoslov. Chem. Commun.* **4**, 335 (1932).
- (112) ŠLENDÝK, I., AND HERASYMENKO, P.: *Z. physik. Chem.* **A162**, 223 (1932).
- (113) STACKELBERG, M., ANTWEILER, H. J., AND KIESELBACH, L.: *Z. Elektrochem.* **44**, 663 (1938).

- (114) STOUT, P. R., LEVY, J., AND WILLIAMS, L. C.: Collection Czechoslov. Chem. Commun. **10**, 129, 136 (1938).
- (115) TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ., No. **42**, Chem. Ser. No. **22**, 65 pp. (1938); Chem. Abstracts **32**, 5685 (1938).
- (116) TAYLOR, H. S.: A Treatise on Physical Chemistry, Vol. 2. D. Van Nostrand Co., New York (1930).
- (117) THANHEISER, G., AND MAASSEN, G.: Mitt. Kaiser-Wilhelm Inst. Eisenforsch. Düsseldorf **19**, 27 (1937).
- (118) TOKUOKA, M.: Collection Czechoslov. Chem. Commun. **4**, 444 (1932); **6**, 339 (1934).
- (119) TOMĚŠ, J.: Collection Czechoslov. Chem. Commun. **9**, 12, 81, 150 (1937).
- (120) TRAPP, C.: Klin. Wochschr. **16**, 374 (1937).
- (121) UHL, F. A.: Z. anal. Chem. **110**, 102 (1937).
- (122) VARASOVA, E.: Collection Czechoslov. Chem. Commun. **2**, 8 (1930).
- (123) VERWEY, E. J. W.: Chem. Rev. **16**, 363 (1935).
- (124) VÍTEK, V.: Collection Czechoslov. Chem. Commun. **7**, 537 (1935); Chimie & industrie **29**, 215 (1933).
- (125) VOPIČKA, E.: Collection Czechoslov. Chem. Commun. **8**, 349 (1936).
- (126) WINKEL, A., AND PROSKE, G.: Ber. **69**, 693, 1917 (1936).
- (127) WINKEL, A., AND PROSKE, G.: Angew. Chem. **50**, 18 (1937).
- (128) WINKEL, A., AND SIEBERT, H.: Z. Elektrochem. **44**, 402 (1938).

OXIDATION AND REDUCTION OF ORGANIC COMPOUNDS AT THE DROPPING MERCURY ELECTRODE AND THE APPLICATION OF HEYROVSKÝ'S POLAROGRAPHIC METHOD IN ORGANIC CHEMISTRY¹

OTTO H. MÜLLER*

Department of Chemistry, Stanford University, California

Received December 18, 1938

CONTENTS

| | |
|--------------------------------------------------------------------------|-----|
| I. Introduction | 95 |
| II. Evaluation of curves | 97 |
| 1. Potential of the non-polarizable (quiet) electrode | 97 |
| 2. Potential of the dropping mercury electrode | 98 |
| 3. Corrections for the iR drop of potential in the solution | 100 |
| 4. Current considerations | 100 |
| III. Reversible oxidation-reduction systems | 101 |
| 1. E'_0 and half-wave potentials in well-buffered solutions | 101 |
| 2. E'_0 and half-wave potentials in unbuffered solutions | 103 |
| 3. Applications | 104 |
| IV. Irreversible reduction | 105 |
| 1. Polarographic apparent reduction potentials | 105 |
| 2. Effect of buffers | 107 |
| 3. Applications | 109 |
| V. Organometallic complexes | 111 |
| VI. Organic compounds that have been studied polarographically | 112 |
| VII. References | 120 |

I. INTRODUCTION

Considerable effort has been made for many years to apply the usually simple and accurate methods of electrochemistry to the study and analysis of organic compounds. Valuable results have been obtained wherever these methods have been applicable. When J. Heyrovský demonstrated the usefulness of the dropping mercury cathode for the study of inorganic reductions, interest in the application of this technic to the field of organic chemistry was aroused. The first organic substance to be reduced at the

* Presented in part before the Division of Physical and Inorganic Chemistry at the Ninety-fifth Meeting of the American Chemical Society, held at Dallas, Texas, April 18-22, 1938.

¹ Upjohn Company Research Fellow, Stanford University, 1937-38. Present Address: Cornell University Medical College, New York City.

dropping mercury cathode was nitrobenzene. This reduction was studied in 1925 by M. Shikata, who was working at that time in the Prague laboratory (90). Plotting current against applied voltage, he obtained a curve which resembled that due to a reduction of metallic ions. After the development of the polarograph (39), many reducible organic compounds were studied; the results were reproducible, the reduction potentials of the various substances were different, and the current was observed to be a function of concentration. Therefore it seemed reasonable to expect that in organic chemistry qualitative and quantitative analyses should be possible by means of the polarographic method. However, the difficulties encountered were far greater than in inorganic reduction; consequently analyses of organic materials were achieved only in isolated cases under well-controlled conditions. At present the qualitative and quantitative analysis of a mixture of organic compounds in unknown solutions is of doubtful value.

In addition to qualitative and quantitative analyses, the polarographic method may be used in studies of the influence of substituted groups and of conjugated double bonds on the reduction potential, in examinations of organic reactions such as complex formation (including catalysis), tautomerism, and polymerization, and in investigations of the nature of the processes at the electrode surface.

The pioneer studies of organic compounds with the polarograph, done in the laboratories of Heyrovský, Shikata, and Semerano, were devoted to the accumulation of sufficient empirical data to serve as the basis for the development of more refined investigations. While interpretations of these experiments had to remain highly speculative until the value of the "half-wave potential" had been demonstrated by Heyrovský and Ilkovič (38), they often gave enough information to permit predictions of the reducibility of compounds. Refinement of the technic and a better understanding of the principles involved have made necessary a number of corrections of these pioneer exploratory investigations. Today almost one half of over four hundred papers on polarography deal with organic reactions, and it appears that the interest in this field is going to exceed that given to inorganic reductions.

A description of the polarographic method has been omitted in this review because it is given in the preceding paper by Kolthoff and Lingane (49), who also discuss the papers dealing with the suppression of maxima and with the effect of the solvent. Different solvents have been used, but only aqueous solutions are considered in this paper.

In the absence of a theoretical discussion of the organic reactions in general treatments of polarography (36, 43, 75), this paper deals with some of the underlying theories and their relation to established principles of electrochemistry. An attempt is made to show the difference

between reversible and irreversible reactions at the dropping mercury electrode and to demonstrate the importance of buffering and the significance of the half-wave potentials.

II. EVALUATION OF CURVES

In the paper (49) preceding this one it has been shown that the reduction of inorganic ions at the dropping mercury electrode is well enough understood to lend itself to accurate and quantitative description. A similar treatment for organic reactions is not yet possible, since they are much more complex because of the greater number of variable factors. Many of these factors, which also apply to inorganic reactions, have been discussed by Kolthoff and Lingane (49). However, the fundamental difference between inorganic and organic reactions is that hydrogen in some form is always involved in the latter, so that the pH and buffer capacity of the solution and the dissociation constants of reductant and oxidant have to be considered. For studies of tautomerism, polymerization, and reaction kinetics it is necessary to have precise knowledge of the temperature, the age of the solution, and the concentration. So far these factors have not received sufficient attention and often have been entirely neglected.

The greatest difficulty in correlating the results of different authors arises from the varying methods of potential measurement. Considering that a single substance may have any number of reduction potential values depending on the pH (see figure 5), a uniform definition and accurate determination of these values becomes of much greater importance than in the inorganic field. It is by no means sufficient simply to give the polarographically *applied voltage* at a certain point of the curve; the anode potential and the drop in potential (iR drop) in the solution have to be known accurately.³ From these three values the real *potential* at the dropping mercury electrode can be calculated and referred to some common standard.

1. Potential of the non-polarizable (quiet) electrode

The voltage applied to the dropping mercury electrode is measured against the potential of the non-polarizable (quiet) electrode or reference

³ This cannot be emphasized strongly enough because even in a recent paper (1) a complete failure to realize its importance has led to some obviously wrong conclusions. The potential of the large mercury layer anode was assumed to be of about the same order in an ammonium chloride and in an alkaline solution and only the "potential across the cell" was considered. On this basis it was concluded that variation in the electrolyte in the solution of an organic compound exerted almost no influence on the polarographic reduction potential of the compound. Such a statement seems to me almost equal to saying that the potential of a quinhydrone electrode is the same at pH 1 and at pH 7, when the reference electrode is a hydrogen electrode.

electrode (58). It is essential that the potential of the reference electrode be accurately known or measurable and that it remain constant even when fairly large currents are drawn, if these polarographic measurements are to have any meaning. Normally this electrode serves as anode, because most polarographic investigations deal with reductions at the dropping mercury cathode. While the anode may consist of any system meeting the above requirements, it has been the custom to use a large mercury layer at the bottom of the electrolytic vessel covered with a solution of a known chloride-ion concentration. Such an electrode has been found very satisfactory in inorganic reductions, as it immediately brings back into solution the metal which was deposited at the cathode, thus insuring an unchanged solution. This particular advantage does not exist for organic reactions, most of which are irreversible. The disadvantages encountered are as follows: (1) The anode potential cannot be assumed to be that of a calomel electrode at the corresponding concentration of chloride ions; for accurate work this potential has to be measured in each solution against a standard half-cell. (2) The potential of the large mercury layer, when used as the cathode, is very poorly poised unless a layer of calomel has previously been deposited on its surface.

A standard calomel half-cell with a large surface, the potential of which is accurately known and which can be checked from time to time, has also been found very satisfactory (59). Its potential remains constant within a few millivolts up to fairly high currents whether it serves as anode or cathode. It is connected to the solution by means of agar bridges. A comparison of the polarograms is simplified because all curves are automatically referred to the same electrode. If potassium chloride-agar bridges are used, the potential measurements are limited to a range between the deposition of chloride and potassium ions. This can be increased, when more positive potentials have to be measured, by using potassium nitrate bridges (58). Perhaps lithium nitrate or tetramethylammonium nitrate bridges would be still better. The disadvantages of this sort of a reference electrode are that the use of the agar bridges introduces some errors due to liquid junction potentials and increases the resistance in the circuit. Repeated resistance measurements become necessary, therefore, when bridges are changed and relatively large currents are drawn, to correct the observed potentials for the iR drop.

2. Potential of the dropping mercury electrode

The potential of the dropping mercury electrode is determined from the polarogram. The applied voltage is read at a well-defined point on the current-voltage curve, then iR is subtracted, and this corrected voltage is added to the known potential of the large reference electrode. The latter

may be referred to any standard. The finding of a suitable point on the current voltage curve for characterization of the reducible substance has been one of the major problems in polarography. It was solved by Heyrovský and Ilkovič (38), who proposed the so-called "half-wave potential". This is the point of inflection of the smooth "S" curve which polarographers have called "wave", and it is very easily measured by drawing two tangents to the ends of the "wave". The tangents must be parallel. A third line midway between the tangents will cut the wave at its point of inflection and give the uncorrected half-wave potential. If the two tangents should not be parallel, one tangent to the most easily determinable end of the wave will suffice. The second line is then drawn parallel to the first at a point which seems to indicate the end of the wave. The half-wave potential is obtained as above; it will be changed but little by slight variations of the second line.

This half-wave potential has been found very convenient in inorganic work (for a more detailed discussion see reference 49), as it is a constant and is independent of the drop time, the rate of flow of mercury, the concentration of the reduced material, and the sensitivity of the galvanometer. Its significance in organic reductions and oxidations has been demonstrated by Muller and Baumberger (59) for reversible oxidation-reduction systems. Irreversible reductions can also be well characterized by it, as will be shown later in this paper. It seems, therefore, justifiable to recommend strongly the use of the half-wave potential for future determinations wherever possible. Its use is still somewhat doubtful for curves with maxima (see, however, references 18 and 124), and its application becomes impossible when the galvanometer deflections are so great that the whole wave cannot be brought onto the polarogram. There is also some indication (1) that the potentials are not constant when more than one organic substance is reduced.

For cases in which the half-wave potential method fails and for an understanding of the older literature the following alternative methods are listed: (1) A 45° tangent is drawn to the curve, and the applied E.M.F. corresponding to the point of contact is used (Heyrovský's old method). (2) The applied E.M.F. at the point of maximum curvature (tangent at $35^\circ 16'$) of the current-voltage curve is used (Semerano). (3) That value of the applied E.M.F. is taken at which an increase of 10 millivolts has caused a rise in current of 1.9×10^{-8} amperes (Shikata).

The most outstanding shortcomings of these methods are that the potentials change with the concentration of the reducible substance, with the sensitivity of the galvanometer, and with the drop time and the rate of flow of the mercury. The exact conditions for standardization of the values always had to be stated in order to make a reproduction of the

results possible. Furthermore, the third method necessitates a high sensitivity of the galvanometer, which makes the determination of a second reduction potential on the same polarogram very difficult. In addition to the variations and combinations of these methods, proposed by different authors, LeBlanc's well-known method of extrapolation to zero current has been used (126, 127); the measurement of R is thus eliminated (see the interrupted lines in figure 3), but the fact that only a very small portion of the wave resembles a straight line makes the extrapolation uncertain, and difficulties arise when more than one wave is present on the polarogram.

If sufficient care is taken the half-wave potentials can easily be obtained with an accuracy of ± 10 millivolts. It should be possible to increase this accuracy to about ± 1 millivolt with the apparatus which is available at present.

3. Corrections for the "iR" drop of potential in the solution

These corrections become necessary when the product of current and resistance exceeds 1 millivolt. This means that in the case of a resistance of 5000 ohms and a galvanometer sensitivity of 2×10^{-9} amperes per millimeter per meter, corrections will have to be made only when the deflection exceeds 1 cm. at a galvanometer sensitivity of 1/10. At a sensitivity of 1/100, the correction will be 10 millivolts if the deflection is 1 cm. and 20 millivolts if the deflection is 2 cm. R is measured by means of a Wheatstone bridge or by determination of the slope of a maximum (44). A simple graphic correction of the half-wave potential for iR is demonstrated by Müller and Baumberger (59) for reversible oxidation-reduction systems. Another method for the correction of the half-wave potential for iR in irreversible reductions is demonstrated in figure 4. Without change of the bridge, indifferent electrolyte, or galvanometer sensitivity, polarograms are taken of the reducible substance at slightly different concentrations. A line drawn through the different half-wave potentials thus obtained will cross the galvanometer zero line at a point corresponding to the potential value which is already corrected for iR .

4. Current considerations

Since the various currents observed have been discussed in detail by Kolthoff and Lingane (49), little need be said about them here. In general the same considerations hold for organic reactions as for inorganic reductions.

The currents are usually plotted on the vertical axis of the polarograms, while the applied voltage is plotted horizontally. If no current is flowing while the potential across the cell is increased, the galvanometer will

remain at rest and a horizontal line on the polarogram will be obtained, which may be called the "galvanometer zero" line. If a substance is reduced and the galvanometer deflection is above the line, the dropping mercury electrode is the cathode; conversely, oxidation of a substance at the dropping mercury electrode, now anode, results in deflection of the galvanometer below this line (59). The currents are cathodic and anodic, respectively.

In inorganic reactions only cathodic currents have been observed, with the exception of anodic currents due to the oxidation of mercury. However, in reversible organic reactions curves are often obtained which go continuously from anodic to cathodic current. The later models of polarographs are equipped with an attachment (55) which makes such continuous anodic and cathodic polarizations possible.

III. REVERSIBLE OXIDATION REDUCTION SYSTEMS

1. E'_0 and half-wave potentials in well-buffered solutions

Certain compounds, when in solution with their reduction products, form systems which are strictly reversible under ordinary conditions. When equilibrium is established, an indicator electrode (58) immersed in such a solution has a definite and reproducible potential, which is a logarithmic function of pH and of the ratio of the concentrations of the oxidized to the reduced substances. These systems are the ones in which electrochemists have been interested, because the potentials obtained can be used for a very rigorous and accurate treatment of oxidation and reduction processes. The classical example is the quinhydrone electrode; oxidant (quinone) and reductant (hydroquinone) are present in equal concentrations, and the potential is entirely a function of pH. The excellent work of Clark and his collaborators on a series of dyes (20) has greatly increased our knowledge of such processes.

Probably because these equilibrium studies seemed quite different from those with a moving electrode surface and continuous reaction, polarographers paid little attention to such systems until 1937, when Müller and Baumberger (59) made a detailed study of the behavior of quinhydrone at the dropping mercury electrode. They used well-buffered solutions containing either quinone, or hydroquinone, or both (quinhydrone). The observed current voltage curves showed that the oxidation of hydroquinone gave rise to a wave of *anodic* current which was identical in appearance with the wave of *cathodic* current due to the reduction of quinone. Quinhydrone in solution dissociates into equivalent amounts of quinone and hydroquinone, and the polarogram shows a wave, half of which is made of anodic current and half of cathodic current (figure 1). At the midpoint of this curve no current flows. This point is the half-wave

potential and corresponds to the well-known E'_0 value that has been obtained with indicator electrodes when the galvanometer serves only as null-point instrument. Interestingly enough, when *only* reductant or *only* oxidant is present, the half-wave potentials obtained (corrected for iR) are identical with the E'_0 value (figure 1). This means that at the half-wave potential the conditions at the electrode surface must have been the same in the three instances, i.e., the concentrations of oxidant and reductant were equal, and the pH was constant. Thus one-half of all the molecules of quinone which diffuse to the cathode in unit time must have been instantly reduced at the half-wave potential, and the hydroquinone formed must have remained at the electrode surface long enough to pro-

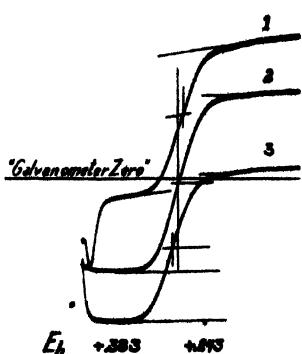


FIG. 1

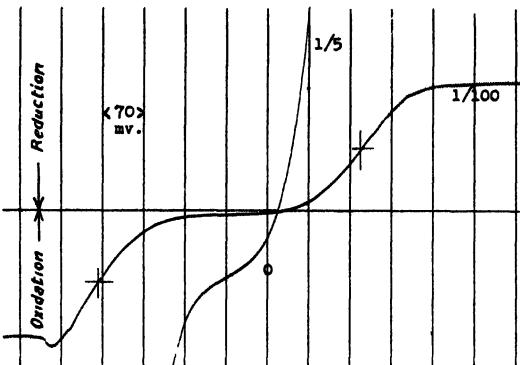


FIG. 2

FIG. 1. Polarogram of (1) quinone reduction, cathodic current, sensitivity 1/70, (2) quinhydrone, anodic and cathodic current, sensitivity 1/70, and (3) hydroquinone oxidation, anodic current, sensitivity 1/40. Buffer solution at pH 6.67 Half-wave potential corrections are indicated.

FIG. 2. Polarogram of quinhydrone ($10^{-2} M$) in 0.1 *N* potassium nitrate solution, unbuffered. 0 indicates zero applied voltage (anode potential = +0.250 volt).

duce the potential. Similarly, one-half of all the hydroquinone ions which diffuse to the anode in unit time are oxidized to quinone and remain at the electrode surface long enough to establish the condition of E'_0 .

For further discussion it must be clear that an electrode can be polarized to any desired potential without any appreciable flow of current as long as no depolarizer is present. Any material which can furnish or accept electrons can act as a depolarizer, but only when the electrode has been brought to a suitable potential. When that potential is reached by applying voltage from some other source, it is kept constant by the transfer of electrons between depolarizer and electrode, that is, by the flow of current. An increase in applied voltage will produce an increase in

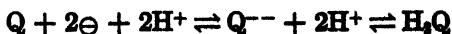
current, while the electrode potential will still be practically the same (59), provided no concentration polarization occurs. This is another way of stating the formula which has been emphasized:

$$\text{Electrode potential} = \text{applied voltage} - (\text{current} \times \text{resistance})$$

The ability to depolarize is, of course, a function of concentration (among other factors), so that an applied voltage can be reached at which polarization of the electrode will be again possible (diffusion current). In inorganic reductions usually a metallic ion is the electron acceptor, and the reduction product is a metal (amalgam). In organic reductions, either the organic molecule or its ion is the acceptor, and the reduction product is a negatively charged organic ion which may or may not undergo secondary reactions. If the reduction product is stable, its dissociation constant is, of course, of prime importance. For instance, if the formed ion is in a medium which is on the acid side of its pK_a , it will at once combine with hydrogen ions and thus alter the pH at the electrode surface unless it is well buffered.

2. E'_0 and half-wave potentials in unbuffered solutions

As is well known, the quinhydrone electrode is used for the determination of pH because its potential is a linear function of pH at constant temperature in the range of pH 1 to pH 8. This dependence is due to the weak acidity of hydroquinone and to the fact that in the electrode reaction only the ionic form of hydroquinone can be considered:⁴



The identity of the half-wave potentials of quinone reduction and hydroquinone oxidation demonstrates that the equilibria shown in this formula must be established with extreme rapidity. The constancy of the pH at the electrode surface implies, furthermore, a similarly rapid dissociation or association of the buffer molecules and ions present.

This last consideration led the author to a polarographic investigation of quinhydrone in unbuffered solutions (57). It was found that the smooth single wave obtained with quinhydrone in buffered solutions (figure 1) broke up into two entirely separate waves when a potassium nitrate solution was used (figure 2). It could be ascertained that these were caused by the reduction of quinone and the oxidation of hydroquinone, respectively. The half-wave potentials, of course, were no longer equal and did not indicate the pH of the potassium nitrate solution, which was practically neutral. The pH calculated from the quinone reduction half-wave potential was about 10, and that from the hydroquinone oxidation half-wave

⁴ For an excellent treatise on this subject see Clark (30).

potential was about 3. These phenomena could be satisfactorily explained by a consideration of the above formula, where hydrogen ion can come only from water or hydroquinone.

In the oxidation of hydroquinone to quinone, two hydrogen ions will be liberated for each hydroquinone molecule oxidized. In an unbuffered solution those hydrogen ions will remain at the electrode surface, changing the reduction potential by a change of pH. Therefore, when one-half of all the hydroquinone molecules are oxidized, the hydrogen-ion concentration at the electrode surface will be equimolar to the original concentration of hydroquinone or quinhydrone. In the above case the quinhydrone concentration was $10^{-8} M$; the pH at the half-wave potential should then be 3, which is the observed value. When the concentration of quinhydrone was changed to $10^{-4} M$, the half-wave potential of the corresponding wave indicated a pH of 4 at the electrode surface.

In the reduction of quinone hydrogen ions are necessary to combine with the hydroquinone ions which are formed. In the absence of a buffer they must come from water, the pH of which is of course changed simultaneously. Around pH 10, however, the dissociation constants of hydroquinone are reached, hydroquinone is now capable of existing in the ionic form in the solution, and, as a rough approximation, no more hydrogen ions are necessary for further reduction. It can be argued from this that the half-wave potential reaches a limit in the neighborhood of the dissociation constant of the reductant in unbuffered solutions.⁵

When to such unbuffered solutions a suitable buffer is added in a concentration below that of the quinhydrone, a third wave appears between the two waves mentioned; this indicates the buffer action at the electrode surface (57). In other words, the trace of buffer added has a certain capacity, limited by its concentration, to keep the pH constant; when the buffer at the electrode surface is exhausted, the pH can change. The sum of these three waves is, of course, a constant and depends on the concentration of quinhydrone. When the buffer concentration is made greater than that of quinhydrone, it is fully able to keep the pH at the electrode surface constant even though all of the hydroquinone be oxidized, or all of the quinone be reduced. In this case the continuous anodic-cathodic curve characteristic of reversible oxidation-reduction systems is obtained.

3. Applications

The polarographic method can be applied to the study of reversible oxidation-reduction systems, the potentials of which fall within the range

⁵ For convenience in this treatment, the two ionization constants of hydroquinone, which are 1.75×10^{-10} and 4×10^{-11} , have been taken as equal. The conclusions reached are not thereby invalidated.

of $E_h = + 0.65$ to $- 1.60$ volts (58, 59, 117). The reversibility of a system can be ascertained when curves of anodic current due to oxidation of the reductant show the same half-wave potential as curves of cathodic current due to the reduction of the oxidant. It has also been demonstrated that semiquinones can be studied polarographically (59). However, because the polarographic method is considerably less accurate than the standard equilibrium procedures with a platinum electrode, its use will probably be limited to exploratory investigations.

IV. IRREVERSIBLE REDUCTION

1. *Polarographic apparent reduction potentials*

The large majority of organic reductions which have been investigated polarographically are not of the type mentioned in the preceding section but are more or less *irreversible*. In many of these reductions a smooth "S" curve is obtained on the polarogram, and a regular shift of the reduction potential with pH is observed; this suggests a reversible process. For this reason they have often erroneously been called reversible, although the corresponding oxidations of the end products have not been possible at the dropping mercury electrode.⁶ The simplest explanation of this phenomenon is the assumption of a reversible step in the reduction, which is on the whole irreversible. Fortunately, analogies in the electrochemical literature were available to test this supposition, which was first made by Müller and Baumberger (60). Conant (21) had occasion to study a number of irreversible reductions which proceeded at different rates depending on the oxidation-reduction potential of the system which was used for the reduction. He concluded that the first step in the reduction must have been reversible and instantaneous, while the next, irreversible step was slow enough to permit measurement. The irreversible system could be characterized by a potential, the "apparent reduction potential" (A.R.P.). This was defined by Conant as the potential of a "critical reagent" which would just cause "appreciable reduction" (20 to 30 per cent in 30 min.).

Only a few of the systems described by Conant have been studied polarographically. Müller and Baumberger (60) evaluated some data in the polarographic literature for dinitrobenzene, nitrobenzene, and maleic acid, and found fair agreement with the A.R.P. of Conant (21). As a further check, the two polarograms in figures 3 and 4 were prepared in this laboratory.⁷ The half-wave potentials are independent of concentration in

⁶ The nature of the end product is of course never known with certainty, because not enough is formed at the mercury droplets to permit an analysis. A reasonable guess can, however, be made on the basis of long-time experiments with larger electrodes.

⁷ Unpublished results.

these highly acid solutions and can be measured easily. The corresponding A.R.P.'s of Conant are indicated by arrows. It may be seen that they differ from the half-wave potentials, but that they coincide with the

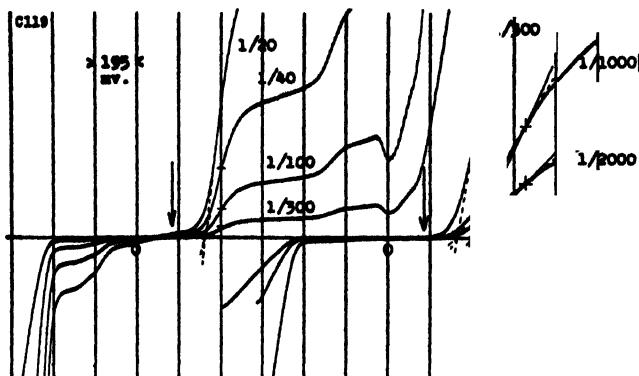


FIG. 3. Polarogram of nitrobenzene in acetone and nitric acid. Comparison of half-wave potentials and Conant's (21) A.R.P. (↓). At sensitivities 1/1000 and 1/2000 the correction for iR is drawn in. Notice also LeBlanc's method of extrapolation to zero current (interrupted lines); iR is graphically eliminated.

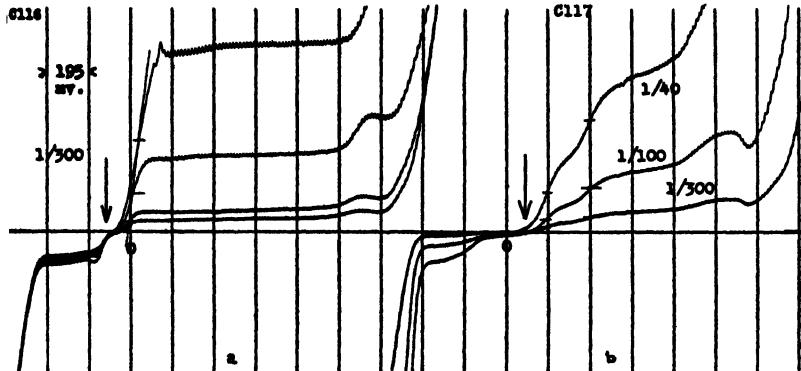


FIG. 4. Polarograms of (a) azobenzene in acetone and nitric acid and (b) dinitrobenzene in acetone and nitric acid. Comparison of half-wave potentials and Conant's (21) A.R.P. (↓). Note graphic correction of the half-wave potentials of azobenzene (see text). 0 indicates zero applied voltage (anode potential = +0.247 volt, referred to the hydrogen electrode). The dropping mercury electrode is polarized negatively to the right of 0 and positively to the left of 0; each abscissa represents 195 millivolts.

beginning of a rise in current due to the onset of reduction. This can be expected, considering the methods of measurement; while Conant measured the A.R.P. indirectly by observing the slow, irreversible process, the

polarograph appears to record selectively the reversible portion of the reaction. We may designate those potentials which are not truly reversible, which show dependence on pH, and which demonstrate partial reversibility by means of smooth "S" curves as "polarographic apparent reduction potentials" (P.A.R.P.).

Consequently, irreversible oxidations of this nature should be called "polarographic apparent oxidation potentials" (P.A.O.P.). Because of the difficulty of measuring very positive potentials at the dropping mercury electrode, those compounds which show an apparent oxidation potential, measured by Conant, cannot be compared. Recently, however, a typical case was found by Kodíček and Wenig (47) in vitamin C, which can be oxidized polarographically but cannot be reduced.

It is premature to speculate on the nature of the reversible step, but it has been suggested that electromotively active ions and molecules may exist which change slowly into an inactive form. Also, it has been suggested that hydrogen ions may be reduced at the electrode to hydrogen atoms which then combine in the nascent state with the organic molecule.

2. Effect of buffers

That buffering is as essential in these reactions as in the reversible ones may be seen from a comparison of some curves of Shikata (91, 100), who was the first to study the influence of pH on the polarographic reduction potentials of organic compounds. In the earlier work (97 to 100) Shikata and Tachi measured the pH of their solutions by means of quinhydrone or hydrogen electrodes, but did not concern themselves with the capacity of the solutions to keep this pH constant. As a result these curves show peculiar inflections around neutrality (see the curves of benzil and diacetyl reproduced in figure 5). When buffer solutions were employed in the later work of Shikata and collaborators (91, 92, 101, 103), relatively smooth curves were obtained (see the curves of nitrobenzene and *o*-dinitrobenzene in figure 5).

It may be worthwhile to restate that pH is an expression of *intensity*; the concentration of the buffer determines the *capacity* of the solutions to keep the pH constant. This has been pointed out repeatedly, but is still occasionally neglected. As a rule one may consider the solution well buffered if the concentration of the buffer is one hundredfold that of the substance which is being reduced or oxidized.

Outside of Shikata's work relatively little has been reported on the change of reduction potentials with a change in pH of the solutions. Examples of potential-pH curves are given in figure 5; they include an aldehyde (benzaldehyde (120)), ketones (benzil, diacetyl (100)), unsaturated acids (fumaric and maleic acids (124)), and aromatic nitro compounds

(nitrobenzene and *o*-dinitrobenzene (91)). Our knowledge of the electrode processes in these reductions is still too meagre to permit satisfactory explanations of the inflections seen in these curves.

Organic reduction potentials measured in neutral potassium chloride or similar unbuffered solutions must be interpreted with caution. As an example of the fallacies encountered in the interpretation of polarograms obtained in unbuffered solutions may be cited the case of acetylacetone. This compound has been reported reducible by several investigators (100,

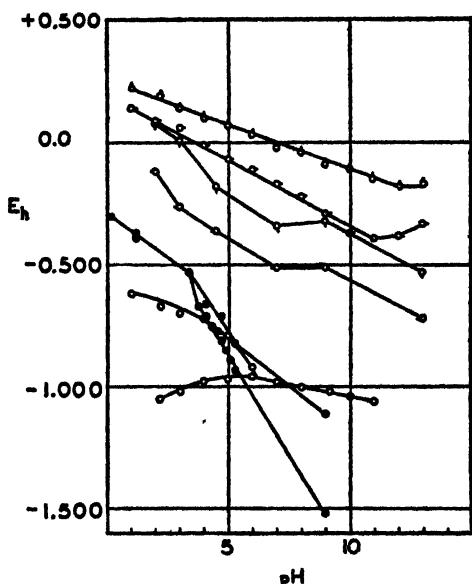


FIG. 5

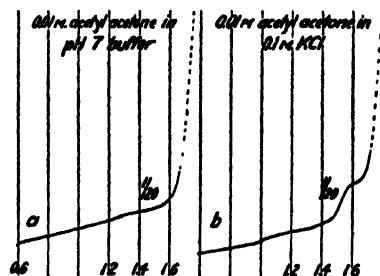


FIG. 6

FIG. 5. Potential-pH curves. δ , *o*-dinitrobenzene; α , nitrobenzene; γ , benzil; β , diacetyl; σ , maleic acid; \varnothing , fumaric acid; τ , benzaldehyde.

FIG. 6. Curves for acetylacetone (a) in a solution buffered at pH 7 and (b) in 0.1 M potassium chloride

127), who worked with unbuffered solutions. A wave is indeed observed when the reduction is carried out in potassium chloride solution (see figure 6b), but the wave disappears when the solution is buffered at pH 7 (figure 6a). Titrations showed⁷ that the unbuffered solution was acid, no doubt owing to the dissociation of the enol form of acetylacetone; therefore it seems probable that the wave was caused by the reduction of free hydrogen ions (see also reference 1).

It has been customary to report reduction potentials of organic compounds obtained in ammonium chloride solutions. This is an unfortunate

choice of electrolyte, because it constitutes but one component of a buffer. As is well known, a buffer is most effective when its two components, acid and salt, are in equal concentration (123). Therefore it seemed probable that when the half-wave potential is reached during a reduction, the pH at the mercury surface would be different from that of the solution. This conclusion was verified using quinhydrone as a criterion (57). Depending on the concentration of ammonium chloride, the pH at the electrode surface, as indicated by the half-wave potential of quinone reduction, was 8, 9.3, 10.0, and 10.2 when the ammonium chloride concentrations were 10^{-1} , 10^{-2} , 10^{-3} , and $10^{-4} M$, respectively. The oxidation of hydroquinone, however, proceeded as if the solution had been unbuffered. At the same time the pH of the solution was found by indicators to be about 5. The results reported in the literature must, therefore, be taken as representing reduction potentials not in a slightly acid solution but rather in a slightly alkaline solution.

Brdička (7, 8, 9) recommends the addition of ammonia to the ammonium chloride solution as a buffer for the cysteine determination. Since Brdička's results are essentially the same as those obtained in phosphate buffers (11), it would be of interest to study this buffer system using quinhydrone.

3. Applications

It is impossible in this short review to do justice to the many excellent papers that deal with applications in this field of irreversible organic reductions. They are especially numerous because obviously the polarographic method excels, owing to the fact that it picks out the reversible step and often gives an indication of the nature of the reaction and the degree of reversibility. Here may be mentioned only examples of the most important and promising types of investigations. Since the conclusions reached in these papers are based on experiments which were carried out under identical conditions, although not always in perfectly buffered solutions, they may have significance. It should be strongly urged, however, that future workers report half-wave potentials which are accurately measured, corrected for iR , and referred to the most common standard in electrochemistry, which is the hydrogen electrode. Furthermore, all measurements should be carried out in well-buffered solutions of known pH. In the table of organic compounds that is included in this discussion, the nature of the study and the corresponding reference are given; the latter is given in italics if the compound has been studied in some detail. Because of the lack of uniformity in the method of measurement, a comparison of the potentials stated in the literature is useless, and a recalculation of the values is impossible because the conditions of measurement are not well enough defined.

In the first studies on the effect of substitutions in an organic molecule upon its reduction potential, Shikata and coworkers investigated different ketones and aromatic nitro compounds (91, 92, 98, 100, 104, 106). On the basis of extended research, Shikata and Tachi formulated the "electronegativity rule of reduction potentials" (102), which states that organic compounds are more easily reduced as more electronegative groups are substituted in the same compound. This is of interest because the dissociation constants (102, 18) and the Raman and absorption spectra (127) of compounds are similarly influenced. The following two lists of Shikata may serve as examples; the compounds are listed in the order of decreasing reducibility:

benzil/diphenyltriketone/diacetyl/dibenzoylmethane/benzoylacetone/
benzoin/benzophenone/acetophenone/acetoin/acetylacetone/acetone

and

dinitrophenol/dinitrobenzene/nitrophenol/nitrobenzene/nitroaniline

Similar studies were made by Semerano and Chisini (81), who used benzaldehyde as a reference. They arranged the following groups in the order in which they change the reduction potential to more negative values:

o-Cl/*m*-Cl/*p*-Cl/*o*-CH₃/H/*m*-CH₃/*p*-CH₃/*o*-OH/CH₃O/*p*-OCH₃/
4-OH, 3-OCH₃

Many more aldehydes and ketones were studied by other investigators, notably Winkel and Proske (126, 127) and Adkins and Cox (1). As far as a comparison of the results of different authors is permissible, it shows that the above influence of different groups upon the reduction potential is essentially correct.

While the aromatic series show such easily reproducible differences in reduction potential, the saturated aliphatic ketones and aldehydes that have been studied are reduced at about the same potential, regardless of the length of the chain. Aldehydes, in general, are more easily reduced than ketones, with the notable exception in the sugar series, where fructose is reducible while glucose is not (40). This fact has been used for analytical purposes (40, 41, 46).

Organic acids in general are not reducible unless they contain carbonyl groups or ethylene linkages. Even in the latter case reduction is possible only when the double bonds are conjugated. This was demonstrated by the investigations of Schwaer (71), who also found that under certain conditions the *cis*- and *trans*-isomers of acids could be distinguished by their polarographic potential (see also 78, 79, 88). As might be expected,

unsaturated aldehydes and ketones are relatively easy to reduce (1). Semerano and Chisini (83) studied cinnamic acid, cinnamaldehyde, and hydrocinnamaldehyde for the purpose of comparing the influence of a C—O group on the reduction potential with that of a C=C linkage.

After it had been established that the currents obtained in the reductions of some organic compounds were proportional to concentration, these currents could be used for the study of reaction kinetics, polymerization, and tautomerism. So far relatively few such experiments have been done, but they indicate the possibilities of polarography.

Herasymenko (34) studied the rate of formation of fumaric acid from molten malic acid by polarographing samples at different time intervals, making use of the fact that malic acid does not produce a wave on the polarogram. Semerano and dePonte (86) found that waves of benzaldehyde reduction decreased with time, owing to the Cannizzaro reaction (see, however, Tokuoka (120)). The inversion of sucrose was followed polarographically by Heyrovský and Smolef, who measured the currents due to the fructose formed (40). Polymerization studies were made on formaldehyde (45) and pyruvic acid (60). Some papers on keto-enol tautomerism (1, 60, 77, 82, 100, 113) have also been published, but verification of the results by comparison with established methods is still lacking. A very clear demonstration of the reliability of polarographic results was made by Borcherdt and Adkins (5). They studied the rate of tautomerization of an optically active azomethine by noting the loss in rotatory power of the reaction mixture and by polarographic analysis of the products of hydrolysis of the components.

As has been pointed out in the introduction, the analysis of several organic compounds in the same solution is possible only under well-standardized conditions. An analysis of a mixture of compounds in an unknown solution is, at present, impossible. For instance, the analysis of pyruvic acid cannot be carried out in blood or urine until satisfactory methods for its separation from the rest of the solution are available. Therefore, one of the most important tasks in the field of organic polarography seems to be to find suitable procedures for the fractionation of solutions in order to make the method available for analytical purposes.

V. ORGANOMETALLIC COMPLEXES

The organometallic complexes may be divided into an inorganic and an organic group, depending on the nature of the electrode reaction.

In the inorganic group a metallic ion is reduced at a potential which varies with the nature of the organic component and with the relative concentrations of the metallic and organic components of the complex. Because the latter may be influenced by hydrogen-ion concentration, a

knowledge of the pH of the solutions is essential. A discussion of this group of complexes is found in the preceding review (49).

In the organic group the metallic component acts only as a catalyst for the reduction of hydrogen. Such catalytic reactions were first observed by R. Brdička in 1933 (7). He found well-reproducible maxima on the polarograms when sulfur-containing proteins or amino acids were added to buffered cobalt salt solutions. His further investigations (8 to 16) show that these reactions can be used for analytical purposes. Sladek and Lipschütz (109) observed that amino acids that do not contain sulfur (arginine, histidine, and others) can also produce specific waves in similar cobalt solutions.

Brdička's work may be summarized as follows: In solutions of ammonium chloride to which ammonia has been added (or in otherwise buffered solutions (11)), cysteine or cystine and sulfur-containing proteins cause characteristic reduction curves when cobalt or nickel salts are present. These curves are much higher (three hundred fold) than those due to the normal reduction of cystine in the absence of metals (8). Amino acids can be distinguished polarographically from the proteins, because low molecular compounds (cystine, cysteine) give the reaction only with divalent cobalt salts, while proteins give similar reactions also with trivalent cobalt salts. The catalytic reduction maxima do not increase proportionally with concentration but reach a maximal height; calibration curves are therefore necessary for quantitative analysis. Cystine is first reduced to cysteine; its reduction curve is twice as high as that of cysteine for an equimolar solution, but does not differ otherwise. A differentiation between cystine and cysteine is possible only by indirect means: The $-SH$ group reacts with monoiodoacetate to form hydrogen iodide and $-SCH_2COO^-$, which do not show the polarographic reaction. In monoiodoacetate the $-S-S-$ group remains unaltered.

According to Brdička (8), these specific curves obtained from solutions containing organic and metallic ions are to be attributed to a catalyzed reduction of hydrogen from the organometallic complex. The abnormal height of the wave (maximum) speaks for a catalytic process (see also reference 56). Further evidence that the reaction is catalytic in nature is provided by the negative influence of arginine, tryptophan, histidine, casein, and other substances (109, 68), which suppress the cysteine reaction completely.

Owing to lack of space, it is impossible to discuss here the many applications which these reactions have found in the field of biochemistry.

VI. ORGANIC COMPOUNDS THAT HAVE BEEN STUDIED POLAROGRAPHICALLY

In table 1 are listed the organic compounds that have been studied polarographically. The literature has been covered up to June, 1938.

TABLE 1
*Organic compounds that have been studied polarographically**

| COMPOUND | | REFERENCES |
|----------------------------------------|-------|-----------------------------------------|
| Acetal..... | - , D | 42 |
| Acetaldehyde..... | R, D | 1, 41, 45, 51, 84, 94, 108, 110, 126 |
| Acetamide..... | - | 127 |
| 4-Acetaminophenylarsonic acid..... | R | 18 |
| Acetic acid..... | - | 30, 65, 94, 108, 109, 126 |
| Acetoacetic ester..... | R, D | 82 |
| Acetoin..... | R | 98, 127 |
| Acetone..... | R | 78, 98 |
| | - | 1, 126 |
| Acetonylacetone..... | R | 100 |
| | - | 127 |
| Acetophenone..... | R, D | 1, 5, 6, 98, 117, 126 |
| Acetylacetone..... | R | 82, 100, 127 |
| | - | 1 |
| Acetylbenzoylmethane..... | R | 1 |
| α -Acetylbutyrophenone..... | R | 1 |
| α -Acetylcoprophenone..... | R | 1 |
| Acetylenedicarboxylic acid..... | R, D | 71 |
| Aconitic acid..... | R, D | 71, 78, 80, 88, 108 |
| Acridine derivative..... | R | 42 |
| Acridone..... | R | 42 |
| Adenine..... | MS | 69 |
| Adipic acid..... | - | 127 |
| Adrenaline..... | MS | 76 |
| Alanine..... | - | 18, 109 |
| Albumin..... | C | 7, 9, 11, 37, 53, 68, 125 |
| Alizarin red..... | MS | 66 |
| Alkaloids..... | R, MS | 63, 27 |
| Allyl alcohol..... | - | 118 |
| p-Aminoasobenzene..... | R, D | 103 |
| 4-Aminophenylarsonic acid..... | R | 18 |
| Anisaldehyde..... | R | 1, 81, 127 |
| Anthraquinone..... | R | 1 |
| L-Arabinose..... | - | 40 |
| Arginine..... | C | 7, 109 |
| Arsphenamine..... | - , D | 10 |
| Asparagine..... | C | 7, 37, 62 |
| Atropine sulfate..... | MS | 27 |
| Aurantia..... | MS | 66 |
| Azobenzene..... | R | 101, 117 |
| Benzal <i>tert</i> -butyl acetone..... | R | 1 |

* - means that the substance is not reducible; C = catalyst (organometallic); D = determinable; MS = maximum suppressor (see 49); R = reducible.

TABLE 1—Continued

| COMPOUND | | REFERENCES |
|-----------------------------------------------|-------|--------------------------------------|
| Benzalacetone..... | R, D | 1 |
| Benzaldehyde..... | R | 1, 81, 86, 94, 108, 120, 126, 127 |
| Bensamide..... | — | 127 |
| Benzil..... | R, D | 1, 100, 117, 126, 127 |
| Benzoic acid..... | R | 97, 120, 126 |
| Benzoin..... | R, D | 1, 77, 98, 117, 127 |
| Benzophenone..... | R, D | 1, 78, 98, 126 |
| Benzopyrone..... | R, D | 1, 54 |
| p-Benzoquinone..... | R, D | 1, 59, 127 |
| Benzoylacetone..... | R, D | 100, 117 |
| Benzoyl- <i>o</i> -hydroxybenzoylmethane..... | R | 1 |
| Benzylbenzoylacetoxymethane..... | R | 1 |
| Benzyl phenyl ketone..... | R | 1 |
| Betaine..... | MS | 69 |
| Biebrich scarlet..... | MS | 66 |
| Bilirubin..... | R, D | 111 |
| Blood pigments..... | | 17 |
| Bromoacetone..... | R | 126 |
| <i>m</i> -Bromobenzaldehyde..... | R | 127 |
| <i>p</i> -Bromobenzaldehyde..... | R | 127 |
| 2-Bromoethyl phenyl ketone..... | R | 1 |
| Bromoisopropyl phenyl ketone..... | R | 1 |
| <i>p</i> -Bromophenyl methyl ketone..... | R | 1 |
| Brucine..... | MS | 63 |
| <i>tert</i> -Butyl benzoylmethyl ketone..... | R | 1 |
| Butyraldehyde..... | R | 1, 83 |
| <i>n</i> -Butyric acid..... | MS | 65 |
| Caffeine..... | —, MS | 62 |
| Camphor..... | R | 107, 115, 116 |
| Caproates..... | MS | 69 |
| Carbazole..... | — | 42 |
| Carbon tetrachloride..... | R | 52 |
| Casein..... | MS | 53, 68 |
| Catalase..... | | 17 |
| Chloroacetone..... | R | 126 |
| <i>p</i> -Chloroacetophenone..... | R, D | 6, 126 |
| <i>m</i> -Chlorobenzaldehyde..... | R | 81, 127 |
| <i>o</i> -Chlorobenzaldehyde..... | R | 81, 127 |
| <i>p</i> -Chlorobenzaldehyde..... | R | 81, 127 |
| <i>p</i> -Chlorobenzophenone..... | R, D | 1, 5, 6 |
| Cholestanone..... | R | 1 |
| Cholesterol..... | — | 109 |
| Choline hydrochloride..... | R | 64 |
| Chromone..... | R, D | 1, 54 |

TABLE I—Continued

| COMPOUND | | REFERENCES |
|------------------------------------------|-------|------------------------------|
| Cinchonidine | R | 63 |
| Cinchonine | R | 63 |
| Cinnamal benzophenone | R | 1 |
| Cinnamic acid | R | 71, 83, 108 |
| Cinnamaldehyde | R | 83, 94, 108 |
| Citraconic acid | R | 71, 78 |
| Citric acid | — | 35, 48, 71 |
| Codeine | R | 63 |
| Coproporphyrin | | 17 |
| Creatine | — | 7, 37, 62 |
| Creatinine | — | 7 |
| Crotonaldehyde | R | 1, 83 |
| Crotonic acid | R | 108 |
| | — | 71 |
| Cyclohexanone | — | 126 |
| Cyclopentanone | — | 1, 126 |
| Cysteine | C | 8, 9, 15 |
| Cysteinylglycine | C | 8 |
| Cystine | C | 7, 8, 9, 15, 23, 67, 122 |
| Desoxybenzoin | R | 77 |
| Diacetyl | R, D | 1, 84, 85, 94, 100, 117, 187 |
| Dibenzalacetone | R | 1 |
| Dibenzoylethane | R | 1, 100 |
| 2,6-Dibromophenol indophenol | R | 59 |
| sym-Dichloroacetone | R | 126 |
| 2,6-Dichlorophenol indophenol | R | 59 |
| 2,4-Dichlorophenylarsonic acid | R | 18 |
| Dicyanogen | R | 19 |
| Diethylbibenzoylmethane | R | 1 |
| Dimethylaminocazobenzene | R | 114 |
| Dimethylquinoxaline | R, D | 59 |
| m-Dinitrobenzene | R, D | 91 |
| o-Dinitrobenzene | R, D | 91 |
| p-Dinitrobenzene | R, D | 91 |
| 2,4-Dinitrophenol (α) | R, D | 92 |
| 2,6-Dinitrophenol (β) | R, D | 92 |
| 2,5-Dinitrophenol (γ) | R, D. | 92 |
| Diphenylamine | — | 42 |
| 1,2-Diphenylethylene glycol | — | 1 |
| Diphenyltriketone | R | 100 |
| Dyes | MS | 66 |
| Dypnone | R | 1 |
| Erythrosin | MS | 66 |
| 2-Ethylchromen-4-ol | D | 54 |

TABLE 1—Continued

| COMPOUND | | REFERENCES |
|---------------------------------|---------|-------------------------------------------------|
| 2-Ethylchromone..... | R | 1, 54 |
| Ethyl isobutyl ketone..... | — | 126 |
| Euglobulin..... | C | 7 |
| Fenchone..... | R | 107 |
| Flavanone..... | R | 1 |
| Flavone..... | R | 1 |
| Flazine..... | R | 42 |
| Formaldehyde..... | R, D | 45, 126 |
| Formamide..... | — | 127 |
| Formic acid..... | —, MS | 30, 65, 126 |
| Fructose..... | R, D | 40, 41, 46 |
| Fuchsin (acid and base)..... | MS | 66, 76 |
| Fumaric acid..... | R, D | 29, 31, 33, 34, 71, 78, 79, 80, 87, 108, 124 |
| Furan..... | — | 121 |
| Furfural..... | R, D | 94, 95, 108, 121 |
| Furfuralacetone..... | R | 1 |
| Furyl acetyl ketone..... | R | 1 |
| Furyl benzoylmethyl ketone..... | R | 1 |
| Fuscainic acid..... | MS | 69 |
| Galactose..... | — | 40 |
| Gelatin..... | — | 7 |
| Globulin..... | C | 7, 125 |
| Glucose..... | — | 40, 94, 109 |
| Glutaminates..... | MS | 69 |
| Glutaric acid..... | — | 127 |
| Glutathione..... | C | 8 |
| Glycerol..... | | 48 |
| Glyceraldehyde..... | R | 127 |
| Glycine..... | —, R, C | 7, 12, 16, 18, 37, 62, 109 |
| Glyoxal..... | R | 127 |
| Helianthine..... | MS | 66 |
| Hematin..... | R | 17, 99 |
| Hemin..... | | 28, 99 |
| Hemoglobin..... | R | 17, 99 |
| Heptaldehyde..... | R | 1 |
| Hexaldehyde..... | R | 1 |
| Histidine..... | —, C | 94, 109 |
| Humic acid..... | —, D | 22, 121 |
| Hydrastine..... | R | 63 |
| Hydrobenzoin..... | — | 83 |
| Hydrocinnamaldehyde..... | R, D | 83 |
| Hydroquinone..... | R | 69 |

TABLE 1—Continued

| COMPOUND | | REFERENCES |
|--------------------------------------------|-------|------------------------------------|
| Hydroxyacetone..... | R | 127 |
| <i>m</i> -Hydroxybenzaldehyde..... | R | 1, 127 |
| <i>p</i> -Hydroxybenzaldehyde..... | R | 127 |
| <i>o</i> -Hydroxyphenyl methyl ketone..... | R | 1 |
| Hymatomelanic acid..... | R | 121 |
| Indanthrene..... | MS | 69 |
| Indole..... | R | 62 |
| Insulin..... | C, D | 7, 61, 122 |
| Iodoacetone..... | R | 126 |
| Isoamyl alcohol..... | — | 95 |
| Isobutyraldehyde..... | R | 1, 95 |
| Isocrotonic acid..... | — | 71 |
| Isopropyl phenyl ketone..... | R | 1 |
| Isovaleraldehyde..... | R | 94, 95 |
| Isovaleric acid..... | MS | 65 |
| Itaconic acid..... | R | 108 |
| | — | 71, 78 |
| Lactic acid..... | — | 48, 94 |
| Lactones..... | R | 40 |
| Lactose..... | — | 40 |
| Lecithin..... | — | 109 |
| Leucine..... | —, MS | 7, 69 |
| Levulic acid..... | — | 71 |
| Lysine..... | — | 109 |
| Lyxose..... | — | 40 |
| Maleic acid..... | R, D | 29, 32, 71, 78, 79, 80, 87, 124 |
| Malic acid..... | D | 34, 41, 71 |
| Malonic acid..... | — | 30, 127 |
| Maltose..... | — | 40 |
| Mannitol..... | — | 48 |
| Mannose..... | — | 40 |
| Melanoidin..... | R | 22 |
| Mesaconic acid..... | R | 71, 106 |
| Mesityl oxide..... | R | 1 |
| Methemoglobin..... | R | 99 |
| <i>p</i> -Methylacetophenone..... | R | 126 |
| Methyl benzyl ketone..... | — | 126 |
| Methylene blue..... | R | 58 |
| | MS | 25, 66 |
| 4-Methyl ether of phenylarsonic acid..... | R | 18 |
| Methyl ethyl ketone..... | R | 98 |
| | — | 126 |

TABLE 1—Continued

| | | REFERENCES |
|-----------------------------------------|----|------------------|
| Methyl green..... | MS | 66 |
| Methyl hexyl ketone..... | — | 126 |
| Methyl α -naphthyl ketone..... | R | 126 |
| Methyl nonyl ketone..... | — | 126 |
| Methyl orange..... | MS | 66 |
| Methyl 1,3-pentadienyl ketone..... | R | 1 |
| 4-Methylphenylarsonic acid..... | R | 18 |
| Michler's ketone..... | R | 126 |
| Morphine hydrochloride..... | R | 63 |
| | MS | 27 |
| Mustard gas..... | C | 19 |
| Myosalvarsan..... | — | 10 |
| | | |
| Narcotine..... | R | 63 |
| Neosalvarsan..... | — | 10 |
| Neutral red..... | R | 117 |
| Nicotine..... | D | 74 |
| Nicotinic acid..... | R | 97 |
| <i>m</i> -Nitroaniline..... | R | 104 |
| <i>o</i> -Nitroaniline..... | R | 104 |
| <i>p</i> -Nitroaniline..... | R | 104 |
| Nitrobenzene..... | R | 90, 91, 116, 119 |
| <i>m</i> -Nitrophenol..... | R | 91, 108 |
| <i>o</i> -Nitrophenol..... | R | 91, 108 |
| <i>p</i> -Nitrophenol..... | R | 91, 108 |
| Nitrophenylarsonic acid..... | R | 18 |
| | | |
| Oleic acid..... | — | 108 |
| Orange II..... | MS | 66 |
| Oxalic acid..... | R | 29, 30, 127 |
| | — | 19 |
| Oxamic acid..... | R | 19 |
| Oxamide..... | R | 127 |
| | | |
| Palmitic acid..... | MS | 65 |
| Papaverine hydrochloride..... | MS | 27 |
| Paraformaldehyde..... | R | 45 |
| Paraldehyde..... | — | 126 |
| Peptone..... | C | 11, 68 |
| Petroleum..... | — | 96 |
| Phenanthraquinone..... | R | 1 |
| Phenethylamine..... | — | 6 |
| Phenolarsonic acid..... | R | 18 |
| Phenol indo-2,6-dibromophenol..... | R | 59 |
| Phenolphthalein..... | MS | 66 |
| β -Phenyl- α -alanine..... | C | 109 |

TABLE 1—Continued

| COMPOUND | | REFERENCES |
|-----------------------------------------------|-------|--------------------------------------------------------|
| β -Phenyl- β -alanine..... | C | 109 |
| Phenylarsonic acid..... | R | 18 |
| Phenyl <i>o</i> -carbethoxyphenyl ketone..... | R | 1 |
| 4-Phenyl-3-ethylbutanedione-2,4..... | R | 1 |
| Phenylethyl phenyl ketone..... | R | 1 |
| Piperonal..... | R | 1, 81, 127 |
| Propionaldehyde..... | R | 1, 94, 126 |
| Propionic acid..... | —, MS | 30, 65 |
| <i>n</i> -Propyl phenyl ketone..... | R | 1 |
| Proteins..... | C | 2, 3, 7, 8, 9, 11, 14, 15, 24, 41, 50, 80, 122, 125 |
| Purine group..... | — | 62 |
| Pyridine..... | R | 95, 96 |
| Pyromucic acid..... | — | 1 |
| γ -Pyrone..... | — | 121 |
| Pyronin..... | R | 1, 64 |
| Pyrrolealdehyde..... | MS | 66 |
| Pyruvic acid..... | R | 4 |
| | R, D | 1, 60, 71, 93, 127 |
| Quinhydrone..... | R, D | 58, 59, 127 |
| Quinidine..... | R | 63 |
| Quinine..... | R | 33, 63, 64, 105 |
| | MS | 27 |
| Quinoline..... | R | 1, 62, 118 |
| <i>p</i> -Quinone..... | R, D | 1, 59, 127 |
| Rhamnose..... | — | 40 |
| Rosinduline GG..... | R | 59 |
| Saccharin..... | R, D | 68 |
| Safranine..... | R | 42 |
| Salicylic acid..... | MS | 26 |
| Salicylic aldehyde..... | R | 1, 81, 127 |
| Sorbic acid..... | R | 71 |
| Sorbose..... | R, D | 40, 41 |
| Stearates..... | MS | 69 |
| Stearic acid..... | MS | 65 |
| Strychnine nitrate..... | R | 63 |
| | MS | 27 |
| Suberic acid..... | — | 127 |
| Succinic acid..... | — | 94, 108, 127 |
| Sucrose..... | —, MS | 40, 48, 69, 70 |
| Tartaric acid..... | —, D | 71 |
| Theobromine..... | — | 62 |

TABLE 1—Concluded

| COMPOUND | | REFERENCES |
|-----------------------------------|-------|----------------|
| Thiazane | — | 13 |
| Thioglycolic acid | C | 8 |
| <i>o</i> -Toluualdehyde | R | 81 |
| <i>p</i> -Toluualdehyde | R | 81 |
| Triacetoneamine | — | 1 |
| Trimethylamine | R | 64 |
| Tropaeolin | MS | 66, 76, 79, 80 |
| Tryptophan | —, C | 109 |
| Tyrosine | MS, — | 7, 69, 109 |
| Ultramarine | MS | 69 |
| Urea | — | 37, 62, 109 |
| Uric acid | —, D | 69 |
| Valeraldehyde | R | 1 |
| <i>n</i> -Valeric acid | MS | 65 |
| Vanillin | R | 1, 81, 127 |
| Vernin | MS | 69 |
| Vitamin C | D | 47 |

A reference number given in italics indicates that the compound has been studied in some detail.

VII. REFERENCES

- (1) ADKINS, H., AND COX, F. W.: J. Am. Chem. Soc. **60**, 1151-9 (1938).
- (2) BABIČKA, J.: Rozpr. II. tř. čes. akad. **41**, 1-13 (1931).
- (3) BABIČKA, J.: Příroda **26**, 62-4 (1933).
- (4) BONINO, G. B., AND SCARAMELLI, G.: Progresso tech. econ. naz. **6**, II, 111, 180 (1933).
- (5) BORCHERDT, G. T., AND ADKINS, H.: J. Am. Chem. Soc. **60**, 3-6 (1938).
- (6) BORCHERDT, G. T., MELOCHE, V. W., AND ADKINS, H.: J. Am. Chem. Soc. **59**, 2171-6 (1937).
- (7) BRDIČKA, R.: Collection Czechoslov. Chem. Commun. **5**, 112-28 (1933).
- (8) BRDIČKA, R.: Collection Czechoslov. Chem. Commun. **5**, 148-64 (1933); Biochem. Z. **272**, 104-12 (1934).
- (9) BRDIČKA, R.: Collection Czechoslov. Chem. Commun. **5**, 238-52 (1933); Mikrochemie **15**, 167-80 (1934).
- (10) BRDIČKA, R.: Časopis Českoslov. Lékárnictva **13**, 51-62 (1933); Collection Czechoslov. Chem. Commun. **7**, 457-66 (1935).
- (11) BRDIČKA, R.: Collection Czechoslov. Chem. Commun. **8**, 366-76 (1936).
- (12) BRDIČKA, R.: J. Gen. Physiol. **19**, 843-66 (1936); **19**, 899-906 (1936).
- (13) BRDIČKA, R.: Collection Czechoslov. Chem. Commun. **9**, 76-80 (1937).
- (14) BRDIČKA, R.: Nature **139**, 330 (1937); **139**, 1020 (1937); Bull. intern. acad. sci. Boheme, No. 4, 38 (1937); Biol. Listy **22**, 39-49 (1937); Compt. rend. soc. biol. **128**, 54 (1938); Acta intern. Verein. Krebsbekämpfung **3**, 18 (1938).
- (15) BRDIČKA, R.: J. chim. phys. **35**, 89-98 (1938).

- (16) BADIČKA, R., AND SCHMIDT, O. L. A.: Univ. Calif. Pub. Physiol. 8, No. 9, 119-28 (1936).
- (17) BADIČKA, R., AND TROPP, C.: Biochem. Z. 289, 301-12 (1937).
- (18) BREYER, B.: Ber. 71, 168-71 (1938).
- (19) BÁBINKA, J., AND HETROVSKÝ, J.: Collection Czechoslov. Chem. Commun. 8, 114-24 (1936).
- (20) CLARK, W. M., AND COWORKERS: Hygienic Laboratory Bull. 151, pp. 1-363 (1928).
- (21) CONANT, J. B.: Chem. Rev. 3, 1 (1926).
- (22) ENDERS, C., AND TEBIS, K.: Wochschr. Brau. 54, 286-7 (1937).
- (23) FRANTA, J., AND GOSMAN, B.: Čsl. Oftalmologie 1, No. 1 (1933).
- (24) GAWALOWSKI, K.: Bull. Dermatol. Syphil. 7, (1932).
- (25) GOSMAN, B. A.: Chimie & industrie, Spacial No., pp. 199-203 (June, 1933).
- (26) GOSMAN, B.: Am. Petroleum Inst. Bull. 10, No. 55, 24 (1930); 11, No. 53, 22 (1930).
- GOSMAN, B., AND HETROVSKÝ, J.: Trans. Electrochem. Soc. 59, 249-71 (1931).
- (27) HAMAMOTO, ELI: Collection Czechoslov. Chem. Commun. 5, 427-35 (1933).
- (28) HAUBOWITZ, F., BADIČKA, R., AND KRAUS, F.: Enzymologia 2, 9-16 (1937).
- (29) HERASYMENKO, P.: Z. Electrochem. 34, 74-8 (1928).
- (30) HERASYMENKO, P.: Z. Elektrochem. 34, 129-36 (1928).
- (31) HERASYMENKO, P.: Ukrains. Khem. Zhur. 4, 439-55 (1929).
- (32) HERASYMENKO, P.: Collection Czechoslov. Chem. Commun. 9, 104-8 (1937).
- (33) HERASYMENKO, P., AND ŠLENDÝK, I.: Collection Czechoslov. Chem. Commun. 6, 204-10 (1934).
- (34) HERASYMENKO, P., AND TYVOŇUK, Z.: Collection Czechoslov. Chem. Commun. 2, 77-82 (1930).
- (35) HETROVSKÝ, J.: Časopis Českoslov. Lékařnictva 14, 285-9 (1934).
- (36) HETROVSKÝ, J.: Physikalische Methoden der analytischen Chemie, edited by W. Böttger. Akademische Verlagsgesellschaft, Leipzig (1936).
- (37) HETROVSKÝ, J., AND BABIČKA, J.: Collection Czechoslov. Chem. Commun. 2, 370-9 (1930); Chem. News 141, 369, 385 (1930).
- (38) HETROVSKÝ, J., AND ILKOVÍČ, D.: Collection Czechoslov. Chem. Commun. 7, 198-214 (1935); Chem. Listy 29, 234-7 (1935).
- (39) HETROVSKÝ, J., AND SHIKATA, M.: Rec. trav. chim. 44, 496 (1925).
- (40) HETROVSKÝ, J., AND SMOLNÁ, I.: Chem. Listy 26, 479-84 (1932); Collection Czechoslov. Chem. Commun. 4, 521-30 (1932).
- (41) HETROVSKÝ, J., SMOLNÁ, I., AND ŠTASTNÝ, J.: Věstník Českoslov. Akad. Zemědělské 8, 490-500 (1930); 9, 599-607 (1933).
- (42) HIGASI, TUNETO: Bull. Inst. Phys. Chem. Research (Tokyo) 15, 1060 (1936); Sci. Papers Inst. Phys. Chem. Research (Tokyo) No. 727, 23, 1-128 (1937).
- (43) HOHN, H.: Anleitung fuer die chemische Laboratoriumspraxis, Band III, edited by E. Zintl. Julius Springer, Berlin (1937).
- (44) ILKOVÍČ, D.: Collection Czechoslov. Chem. Commun. 4, 480 (1932).
- (45) JAHODA, F. G.: Časopis Českoslov. Lékařnictva 14, 225-34 (1934); Collection Czechoslov. Chem. Commun. 7, 415-23 (1935).
- (46) JURS, P. C., AND NOLLER, C. R.: J. Am. Chem. Soc. 58, 1251-5 (1936).
- (47) KODÍČEK, E., AND WANIS, K.: Nature 142, 35 (1938).
- (48) KOMÁRK, K.: Collection Czechoslov. Chem. Commun. 7, 404-14 (1935).
- (49) KOTTHOFF, I. M., AND LINGANE, J. J.: Chem. Rev. 24, 1-94 (1939).
- (50) KOTLJAR, A. M., AND PODROUČEK, V.: Časopis Českoslov. Lékařnictva 18, 123 (1938).

- (51) LHOTSKÝ, A.: Zpr. list. kvařného prům. Brno 8, 51-5 (1936).
- (52) MATHESON, L. A., AND NICHOLS, N.: Trans. Electrochem. Soc. 73, 193 (1938).
- (53) MOHR, W., AND WELLM, J.: Milchw. Forsch. 18, 123-80 (1936).
- (54) MOZINGO, R., AND ADKINS, H.: J. Am. Chem. Soc. 60, 669-75 (1938).
- (55) MÜLLER, O. H.: Collection Czechoslov. Chem. Commun. 6, 289-82 (1934).
- (56) MÜLLER, O. H.: Collection Czechoslov. Chem. Commun. 7, 321-5 (1935).
- (57) MÜLLER, O. H.: Paper presented at the Meeting of the American Chemical Society, Milwaukee, 1938.
- (58) MÜLLER, O. H., AND BAUMBERGER, J. P.: Trans. Electrochem. Soc. 71, 169-80 (1937).
- (59) MÜLLER, O. H., AND BAUMBERGER, J. P.: Trans. Electrochem. Soc. 71, 181-94 (1937).
- (60) MÜLLER, O. H., AND BAUMBERGER, J. P.: J. Am. Chem. Soc., in print.
- (61) MULLI, K., AND WERNER, H.: Deut. med. Wochschr. 63, 1941-3 (1937).
- (62) PECH, J.: Collection Czechoslov. Chem. Commun. 6, 126-36 (1934).
- (63) PECH, J.: Collection Czechoslov. Chem. Commun. 6, 190-203 (1934).
- (64) PODRouŽEK, W.: Rec. trav. chim. 44, 591-9 (1925).
- (65) RASCH, J.: Collection Czechoslov. Chem. Commun. 1, 560-70 (1929); Bull. intern. acad. sci. Boheme 30, 306-16 (1929).
- (66) RATMAN, B.: Collection Czechoslov. Chem. Commun. 3, 314-27 (1931).
- (67) RONCATO, A.: Arch. sci. biol. (Italy) 20, 146-71 (1934).
- (68) ROSENTHAL, H. G.: Mikrochemie 22, 233-41 (1937).
- (69) ŠANDERA, K.: Collection Czechoslov. Chem. Commun. 2, 363-9 (1930).
- (70) ŠANDERA, K., AND ZIMMERMANN, B.: Listy Cukrovar. 47, 377-82 (1929); Z. Zuckerind. Čechoslovak. Rep. 53, 373 (1929).
- (71) SCHWAER, L.: Collection Czechoslov. Chem. Commun. 7, 326-35 (1935); Chem. Listy 28, No. 20, 485-9 (1932).
- (72) SCHWEITZER, H., AND LAQUEUR, E.: Rec. trav. chim. 55, 959-62 (1936).
- (73) SEMERANO, G.: Gazz. chim. ital. 62, 959-91 (1932).
- (74) SEMERANO, G.: Giorn. chim. ind. applicata 14, 608-14 (1932).
- (75) SEMERANO, G.: Il polarografo, sua teoria e applicazioni, 2nd edition, 243 pp. Draghi, Padova (1933).
- (76) SEMERANO, G.: Arch. sci. biol. (Italy) 20, 329-42 (1934).
- (77) SEMERANO, G.: Gazz. chim. ital. 65, 273-88 (1935).
- (78) SEMERANO, G.: Mikrochemie 24, 10-5 (1938).
- (79) SEMERANO, G., AND BETTINELLI, G.: Gazz. chim. ital. 66, 744-9 (1936).
- (80) SEMERANO, G., AND BETTINELLI, G.: Mem. accad. Italia 8, 243-53 (1937); 8, 255-94 (1937).
- (81) SEMERANO, G., AND CHISINI, A.: Gazz. chim. ital. 68, 802-18 (1938).
- (82) SEMERANO, G., AND CHISINI, A.: Gazz. chim. ital. 68, 504-9 (1938).
- (83) SEMERANO, G., AND CHISINI, A.: Gazz. chim. ital. 68, 510-18 (1938).
- (84) SEMERANO, G., AND POLACSEK, B.: Gazz. chim. ital. 68, 292-5 (1938).
- (85) SEMERANO, G., AND POLACSEK, B.: Magyar Chem. Folyóirat 64, 75 (1938).
- (86) SEMERANO, G., AND DE PONTE, G.: Gazz. chim. ital. 62, 991-9 (1932).
- (87) SEMERANO, G., AND RAO, I. S.: Mikrochemie 23, 9-16 (1937).
- (88) SEMERANO, G., AND SAETORI, L.: Mikrochemie 24, 180-3 (1938).
- (89) SEUBERLING, O.: Klin. Wochschr. 18, 644 (1937).
- (90) SHIKATA, M.: Trans. Faraday Soc. 21, 42-53 (1925); 21, 53-61 (1925); J. Agr. Chem. Soc. Japan 1, 533 (1925).
- (91) SHIKATA, M., AND HORAKI, N.: Mem. Coll. Agr. Kyoto Imp. Univ. 17, 1-19 (1931).

- (92) SHIKATA, M., AND HOSAKI, N.: Mem. Coll. Agr. Kyoto Imp. Univ. 17, 21-55 (1931); J. Agr. Chem. Soc. Japan 8, 1121-8 (1932).
- (93) SHIKATA, M., AND SHOJI, K.: Bull. Agr. Chem. Soc. Japan 4, 96-7 (1928).
- (94) SHIKATA, M., AND SHOJI, K.: Mem. Coll. Agr. Kyoto Imp. Univ. 4, 75-91 (1927); J. Agr. Chem. Soc. Japan 5, 212 (1928).
- (95) SHIKATA, M., AND TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ. 4, 9-19 (1927); Proc. Imp. Acad. (Tokyo) 8, 226-8 (1926); Ber. ges. Physiol. exptl. Pharmakol. 38, 622 (1926).
- (96) SHIKATA, M., AND TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ. 4, 19-33 (1927); Bull. Agr. Chem. Soc. Japan 8, 53 (1927); Chem. News 137, 183-4 (1928).
- (97) SHIKATA, M., AND TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ. 4, 35-48 (1927); Bull. Agr. Chem. Soc. Japan 8, 95-6 (1927); Chem. News 137, 126 (1928).
- (98) SHIKATA, M., AND TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ. 8, 1-19 (1930); Bull. Agr. Chem. Soc. Japan 4, 44-6 (1928).
- (99) SHIKATA, M., AND TACHI, I.: J. Agr. Chem. Soc. Japan 5, 352 (1929).
- (100) SHIKATA, M., AND TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ. 8, 21-46 (1930).
- (101) SHIKATA, M., AND TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ. 17, 45-55 (1931). J. Agr. Chem. Soc. Japan 8, 954-61 (1932); Bull. Agr. Chem. Soc. Japan 8, 75-6 (1932).
- (102) SHIKATA, M., AND TACHI, I.: J. Chem. Soc. Japan 53, 884 (1932); Collection Czechoslov. Chem. Commun. 10, 834 (1938).
- (103) SHIKATA, M., AND TACHI, I.: J. Agr. Chem. Soc. Japan 9, 207-18 (1933); Bull. Agr. Chem. Soc. Japan 8, 154-5 (1932); Mem. Coll. Agr. Kyoto Imp. Univ. 40, 1-10 (1937).
- (104) SHIKATA, M., AND TAGUCHI, E.: Mem. Coll. Agr. Kyoto Imp. Univ. 29, 1-17 (1934); J. Agr. Chem. Soc. Japan 8, 1225-36 (1932).
- (105) SHIKATA, M., AND TAGUCHI, E.: J. Electrochem. Assoc. Japan 8, 224-9 (1934).
- (106) SHIKATA, M., AND WATANABE, M.: J. Agr. Chem. Soc. Japan 4, 924 (1928).
- (107) SHIKATA, M., AND WATANABE, M.: J. Agr. Chem. Soc. Japan 5, 904 (1929).
- (108) SHOJI, K.: Bull. Agr. Chem. Soc. Japan 8, 96-7 (1927); Bull. Inst. Phys. Chem. Research (Tokyo) 9, 69-78 (1930); Bull. Inst. Phys. Chem. Research (Tokyo) 9, 621-8 (1930); J. Agr. Chem. Soc. Japan 8, 1126 (1927).
- (109) SLÁDEK, J., AND LIPSCHÜTZ, M.: Collection Czechoslov. Chem. Commun. 6, 487-97 (1934); Bull. intern. acad. sci. Boheme (December, 1934).
- (110) SMOLÍK, I.: Collection Czechoslov. Chem. Commun. 8, 699-71 (1930); Chem. News 142, 97-105 (1931).
- (111) TACHI, I.: J. Agr. Chem. Soc. Japan 6, 533 (1931).
- (112) TACHI, I.: Mem. Coll. Agr. Kyoto Imp. Univ. 17, 35-44 (1931).
- (113) TACHI, I.: J. Agr. Chem. Soc. Japan 7, 524 (1931).
- (114) TACHI, I.: J. Agr. Chem. Soc. Japan 8, 227-34 (1933); Bull. Agr. Chem. Soc. Japan 8, 155 (1932); Mem. Coll. Agr. Kyoto Imp. Univ. 40, 11-20 (1937).
- (115) TACHI, I.: J. Agr. Chem. Soc. Japan 10, 320-6 (1934).
- (116) TACHI, I.: J. Agr. Chem. Soc. Japan 11, 734-40 (1935).
- (117) TACHI, I.: J. Agr. Chem. Soc. Japan 13, 692-7 (1936); 13, 698-704 (1936).
- (118) TACHI, I., AND KABAI, H.: J. Electrochem. Assoc. Japan 8, 260 (1935).
- (119) TEISINGER, J.: Čas. ľekářstv. čes. 11, 325-30 (1937).
- (120) TOKUOKA, M.: Collection Czechoslov. Chem. Commun. 7, 393-403 (1935).
- (121) TOKUOKA, M., AND RČÍŠKA, J.: Z. Pflanzenernähr. Düngung Bedenk. 26A, 79-88 (1934).

- (122) TROPP, C.: *Klin. Wochschr.* **17**, 465-9, 1141-8 (1938).
- (123) VAN SLYKE, D. D.: *J. Biol. Chem.* **52**, 525 (1922).
- (124) VOPIČKA, E.: Collection Czechoslov. Chem. Commun. **8**, 349-65 (1936).
- (125) WENIG, K., AND JÍROVEC, O.: *Biochem. Z.* **295**, 405-13 (1938).
- (126) WINKEL, A., AND PROSKE, G.: *Ber.* **69**, 693-706 (1936).
- (127) WINKEL, A., AND PROSKE, G.: *Ber.* **69**, 1917-29 (1936).

THE ELECTRODEPOSITION OF HYDROGEN AND DEUTERIUM AT THE DROPPING MERCURY CATHODE¹

J. HEYROVSKÝ

Charles University, Prague, Czechoslovakia

Received December 19, 1958

The high reproducibility of the polarographic current-voltage curves obtained with the dropping mercury cathode allows one to follow the phenomenon of hydrogen overvoltage in great detail. In this way it was established that the potential of a mercury cathode at which hydrogen is being evolved changes with current density and with the concentration of ions in a manner differing from that of a reversible hydrogen electrode. The old definition of overvoltage as the difference between the potential of the metal being examined and that of a reversible hydrogen electrode in the same solution has no precise meaning.

O vervoltage must be defined by its dependence on the current density and must be referred to the standard potential of a non-polarizable electrode (in this work the decinormal calomel electrode). The first empirical equation defining the overvoltage, π , at the dropping mercury cathode was deduced by Heyrovský (9, 10) and Herasymenko (7, 8) in the form:

$$\pi = 2RT/F \log [H^+]_s - 3RT/2F \log i + K$$

where $[H^+]_s$ is the concentration of hydrogen ions in the surface layer at the mercury cathode in the solution of the strong acid in the presence of excess electrolyte (potassium chloride) and i is the current. Since the average areas of the mercury drops are constant, the current density is taken to be proportional to the current. If concentration polarization at the cathode is taken into account,

$$[H^+]_s = k(i_d - i)$$

where i_d is the limiting (or "diffusion") current due to the deposition of hydrogen ions, we obtain, according to Tomeš (17), the overvoltage relationship in the form

$$\pi = 2RT/F \log (i_d - i) - 3RT/2F \log i + K$$

¹ Presented before the Division of Physiological and Inorganic Chemistry at the Ninety-fifth Meeting of the American Chemical Society, held at Dallas, Texas, April 18-22, 1958.

This formula does not hold exactly for very small currents or for very rapid dropping of mercury, in which case the factor $3/2$ decreases towards unity. Moreover it does not hold for a slow rate of dropping, since in this case the factor $3/2$ increases towards 2.

Heyrovský (11), in correlating the observations of Müller, Novák, and Tomeš, corrected the above formula to

$$\pi = 2RT/F \log (i_d - i) - RT/F \log (1 + \omega i)i + K \quad (1)$$

where ω is the "adsorption coefficient" of the freshly formed hydrogen molecules at the cathodic interface.

When the rate of dropping is great, adsorption of the molecules is hindered by the rapid renewal of the mercury surface, and ω is small. Under these conditions ωi is small compared to unity, and equation 1 takes the form

$$\pi = 2RT/F \log (i_d - i) - RT/F \log i + K'$$

This form of the equation also holds for very small currents, since here again ωi is small compared to 1.

For a small dropping speed, ω is so large that 1 is small compared to ωi . The overvoltage relationship then takes the form

$$\pi = 2RT/F \log (i_d - i) - 2RT/F \log i + K''$$

This equation must hold for infinitely slow dropping speed, i.e., for the overvoltage at a steady mercury surface. This has been well established (3, 5, 6).

Mathematical analysis of equation 1 shows that it agrees with the course of the overvoltage current-voltage curve (11), particularly as regards the position of the so-called "half-wave" potential (see figure 1).

It is significant for the mechanism of the deposition of the different hydrogen isotopes that equation 1 holds just as well for the electrodeposition of deuterium from solutions of deuterium chloride or deuterium sulfate in heavy water, where the adsorption coefficient is approximately eight times greater than in light water and the constant K is more negative.

Figure 1 shows a comparison of the polarographic current-voltage curves in light and in heavy water. These curves were obtained under identical conditions of dropping, temperature, concentration, and galvanometer sensitivity. Utmost care was taken to keep the 99.6 per cent heavy water protected from light water. The coefficient of $\log i$, which is 0.102 in light water near the "half-wave" potential, becomes 0.133 in heavy water. In mixtures of light and heavy water the values of this coefficient, b , as well as those of the "adsorption coefficient," ω , vary as given in table 1.

The experimental results in table 1 were obtained in careful polarographic investigations by Novák (14). This same author obtained the important result that the overvoltage in heavy water decreases at high

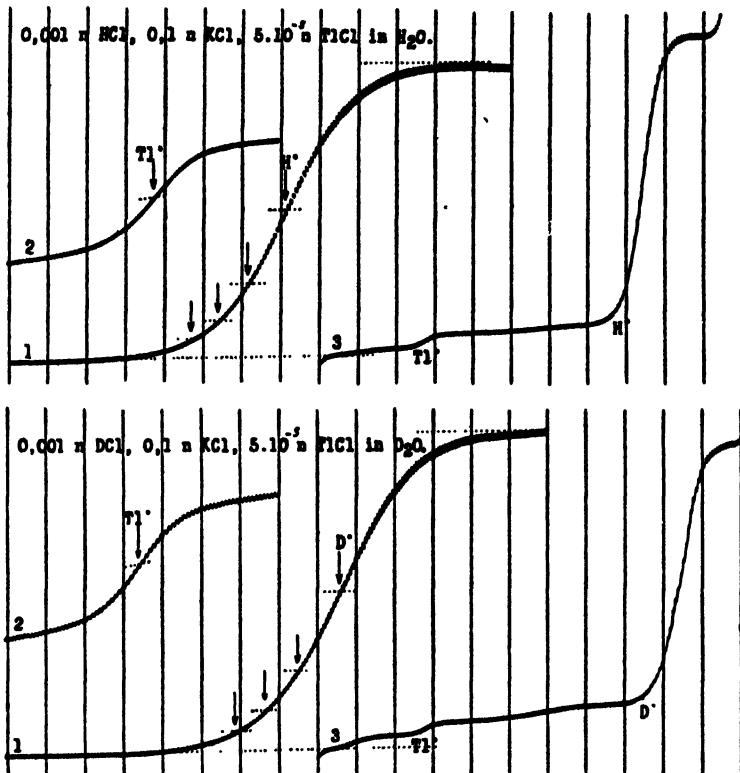


FIG. 1. Comparison of the polarographic current-voltage curves in light and in heavy water. Upper polarogram: 0.001 *N* hydrochloric acid, 0.1 *N* potassium chloride, and 5×10^{-4} *N* thallous chloride in light water. Curve 1: 50 millivolts per complete rotation, from 1.3 volts E.M.F., sensitivity 1/100. Curve 2: 50 millivolts per complete rotation, from 0.35 volt E.M.F., sensitivity 1/2. Curve 3: 200 millivolts per complete rotation, from 0 volt E.M.F., sensitivity 1/100. Lower polarogram: 0.001 *N* deuterium chloride, 0.1 *N* potassium chloride, and 5×10^{-4} *N* thallous chloride in heavy water. Curve 1: same E.M.F. and sensitivity as curve 1 in upper polarogram. Curve 2: same E.M.F. and sensitivity as curve 2 in upper polarogram. Curve 3: same E.M.F. and sensitivity as curve 3 in upper polarogram.

temperature more than in light water. At 60°C. the coefficients, *b*, are in the ratio 0.116/0.111, whereas at 20°C. the ratio is 0.113/0.102. Thus at higher temperatures the overvoltage in heavy water approaches that in light water.

The significant feature of the current-voltage curves in mixtures of light and heavy water is that they show only one homogeneous bend without a trace of double-wave formation (see figure 2). This proves that the isotopic hydrogen ions are deposited at the same rate at the same potential. In other words, the deposition potentials of hydrogen ion and deuterium ion do not differ materially. The difference of 87 millivolts between the evolution of hydrogen from light and from heavy water is large enough to make a double wave quite evident, providing this difference were really that of the deposition potentials. The polarographic curve of a mixture of radium and barium shows a double wave even though the deposition potentials differ by only 60 millivolts (13).²

The fact that there cannot exist different deposition potentials for the hydrogen isotopes refutes any theory which ascribes the greater overvoltage of deuterium to a slower rate of deposition of its ions. The theory

TABLE 1

Differences in overpotentials at the "half-diffusion current" observed with 0.001 N hydrochloric acid, at 20°C., in mixtures of heavy and light water and in pure light water

| POLAROGRAM NO. | D ₂ O | $\pi_1 - \pi'_1$ (OBSERVED) | b | i | $\pi_1 - \pi'_1$ (CALCULATED) |
|----------------|------------------|--------------------------------|-------|-------|----------------------------------|
| volts | | | | | |
| 12 | 0.996 | 0.087 | 0.113 | 8.33 | 0.081 |
| 15 | 0.986 | 0.078 | 0.111 | 7.60 | 0.077 |
| 16 | 0.946 | 0.063 | 0.111 | 4.80 | 0.063 |
| 17 | 0.765 | 0.031 | | 2.50* | 0.035 |
| 18 | 0.498 | 0.015 | 0.104 | 1.41 | 0.017 |
| 33 | 0 | 0 | 0.102 | 1.07 | |

* Value interpolated from the graph.

of Erdey-Grúz and Volmer (5), which assumes the discharge of hydrogen ions to be the slowest process in the cathodic evolution of hydrogen, is disproved by the fact that heavy hydrogen is evolved more slowly than light hydrogen. This theory, originated by Smits (16), postulates a slow rate of deposition of hydrogen ions owing to their dehydration. In heavy water the ions are somewhat less hydrated than in light water, owing to the smaller free energy and smaller dielectric constant of deuterium oxide, and the evolution of deuterium should be facilitated; this is in disagreement with experimental results.

² In the theoretical discussion which follows the reader must distinguish clearly between the process of "deposition (or discharge) of hydrogen ions," which means the cathodic formation of atomic hydrogen in the nascent state, and the evolution of hydrogen, which means the total process of hydrogen molecule formation.

The polarographic results of Novák (14; see also figure 1) show that the deposition of thallous and zinc ions proceeds in heavy water at slightly more positive potentials than in light water. Heyrovský (11) suggests that this is also the case for deuterium ions and for hydrogen ions and that the rate of formation of molecular deuterium from nascent deuterium is

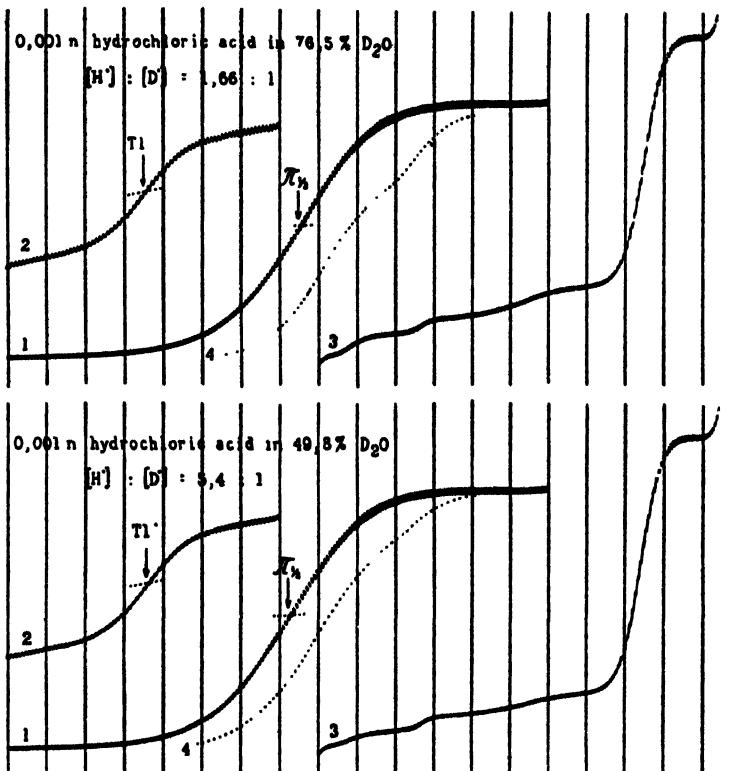


Fig. 2. Polarographic curves of 0.001 N hydrochloric acid. Upper polarogram: in a mixture of 76.5 per cent heavy water and 23.5 per cent light water. Lower polarogram: in a mixture of 49.8 per cent heavy water and 50.2 per cent light water. Curve 1, with sensitivity 3/50 from 1.8 volts E.M.F.; curve 2, with sensitivity 1/2 from 0.35 volt E.M.F.; curve 3, with sensitivity 3/50 from 0 volt; curve 4, expected, if deposition potentials of isotopes were different.

slower than that of hydrogen. Electrolytic hydrogenation which does not include molecule formation (e.g. the electroreduction of fumaric and maleic acids in acid solution) proceeds at the dropping mercury electrode reversibly without overvoltage. A somewhat greater positive potential is obtained in heavy water than in light water for such processes. On the

other hand, the electroreduction of hydrogen peroxide takes place in an alkaline solution in heavy water at a potential which is 87 millivolts more negative than in light water. This must be due to the process of discharge of deuterium ions from alkaline solution and suggests that the ion product constant of deuterium oxide is responsible for the large overvoltage in this case.

The above considerations led Heyrovský (11) to the following theory of overvoltage in light and heavy water and to the deduction of the electrolytic separation coefficients of the hydrogen isotopes.

Hydrogen isotopes are discharged reversibly and at the same potential, forming at the mercury cathode highly active amalgams of atomic hydrogen, potentials of which are given by the following equation:

$$\pi_H = -RT/F \log \frac{P_H}{[H^+]} k = -RT/F \log \frac{P_D}{[D^+]} k = -RT/F \log \frac{P_H + P_D}{[H^+] + [D^+]} \quad (2)$$

in which P_H and P_D denote the activities of atomic hydrogen and atomic deuterium at the cathode. (Drucker (4) has shown recently that the reversible electrode potentials of hydrogen and of deuterium are nearly identical.)

The rate of formation of molecular hydrogen by combination between atoms is very slow on a mercury surface, so that the rate is determined by the reaction (10)



The subsequent neutralization of the particle H_2^+ by an electron at the cathode is supposed to be very rapid. The magnitude of the current is proportional to the product $P_H[H^+]$. In an acid solution of light and heavy water

$$i = k[P_H + P_D] ([H^+] + [D^+])$$

To express the kinetic activity of hydrogen ions in mixtures of light and heavy water, let us use the Grotthuss idea of constant interchange of hydrogen ions with water. Water molecules dissociate at a rate which must be equal to that of the rate of recombination of their ions. The latter is given by the product

$$k'([H^+] + [D^+])([OH^-] + [OD^-])$$

where k' is independent of the composition of the mixture, as it is governed only by the electrostatic properties of the isotopic ions. The above

product is equal to the sum of the four ionic products of light and heavy water, which in the mixture is equal to

$$K_1 C_{\text{H}_2\text{O}} + K'_1 C_{\text{HOD}} + K''_1 C_{\text{HOD}} + K_2 C_{\text{D}_2\text{O}} = a$$

where $C_{\text{H}_2\text{O}}$, C_{HOD} , and $C_{\text{D}_2\text{O}}$ denote mole fractions of the species H_2O , HOD , and D_2O , so that

$$C_{\text{H}_2\text{O}} + C_{\text{HOD}} + C_{\text{D}_2\text{O}} = 1$$

a denotes the rate of dissociation of water molecules in the mixture. Then

$$k''a([H^+] + [D^+])$$

expresses the rate of exchange of hydrogen ions in the mixture, and this is the kinetic reactivity with which they impinge on the cathode interface. Therefore in the kinetic formula for the magnitude of current, $k''a$ must be substituted for k , and

$$i = k''a(P_{\text{H}} + P_{\text{D}})([H^+] + [D^+])$$

Correction for the adsorption of freshly formed hydrogen molecules at the cathodic interface by the theory of Langmuir leads to

$$i = k''a(P_{\text{H}} + P_{\text{D}})([H^+] + [D^+])/(1 + \omega i)$$

Substitution in equation 2 gives

$$\pi = 2RT/F \log ([H^+] + [D^+]) - RT/F \log [i(1 + \omega i)/a] + K$$

which is the general formula of overvoltage for any mixture of light and heavy water.

The difference in overvoltage in pure water and in pure deuterium oxide of the same acidity is accordingly

$$\pi - \pi' = RT'/F \log \frac{a_1}{a_2} \cdot \frac{1 + \omega_2 i}{1 + \omega_1 i} i$$

According to Schwarzenbach, Epprecht, and Erlenmayer (15) $a_1 = 10^{-14}$, $a_2 = 0.185 \times 10^{-14}$, $\omega_1 i$ from the coefficient 0.012 is 1.07, and $\omega_2 i$ from the coefficient 0.113 is 8.33. Hence we calculate

$$\pi - \pi' = RT/F \log \frac{10^{-14}}{0.185 \times 10^{-14}} \cdot \frac{1 + 8.33}{1 + 1.07} = 0.081 \text{ volt}$$

The observed value is 0.087 volt. It should be mentioned that the only results of overvoltage in deuterium oxide published besides those discussed here are the data of Bowden and Kenyon (2). They obtained in 0.1 m deuterium sulfate a difference of 130 millivolts from that in light water, while the coefficient, b , in heavy water = $2RT/F = 0.116$ volt, just as

in light water. The latter value is in perfect agreement with the theory given here for stable electrodes. The difference of 130 millivolts, although much larger than that observed polarographically, is quite compatible with the formula in view of the great increase of adsorption, ω , at a stable mercury electrode.

The agreement between the overvoltage differences calculated and observed in mixtures of heavy and light water, as shown in table 1, is still better. The concentration of HOD had to be calculated. Equations

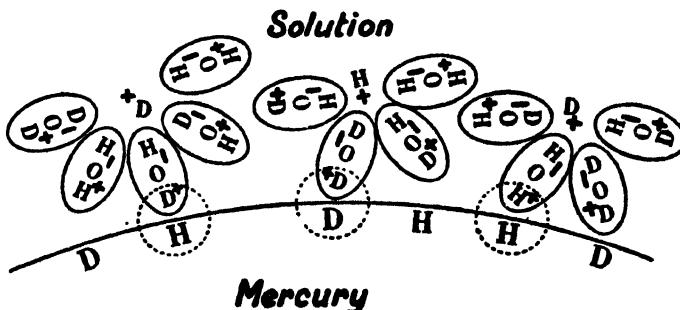


FIG. 3. The slowest step in the formation of molecules of hydrogen evolved at a cathode under overvoltage. The interaction of the electrodeposited hydrogen and deuterium atoms with water molecules in a mixture of light and heavy water to form the particles H_2^+ , HD^+ , and D_2^+ .

worked out by Heyrovský (11, 12) in his original publication on the overvoltage of heavy water give

$$[H^+][OD^-]/C_{HOD} = K'_1 = 0.5 \times 10^{-14}$$

$$[D^+][OH^-]/C_{HOD} = K'_2 = 0.0925 \times 10^{-14}$$

$$K_1 K_2 / K'_1 K'_2 = C_{HOD}/C_{H_2O} \cdot C_{D_2O} = 4$$

$$C_{HOD} = 2x(1-x); C_{H_2O} = (1-x)^2; \text{ and } C_{D_2O} = x^2$$

for any mixture made up of x volumes of heavy water and $(1-x)$ volumes of light water. From the above picture that molecules of hydrogen are formed by the interaction of deposited atoms with hydrogen ions of the solution (see figure 3) the electrolytic separation coefficient of the hydrogen isotope, α ,

$$(H)/(D) = \alpha(1-x)/x$$

may be derived. (H) and (D) denote the numbers of hydrogen and deuterium atoms in the electrolytically evolved hydrogen. Their ratio must be proportional to

$$(2i_{H_2} + i_{HD})/(2i_{D_2} + i_{HD})$$

where i_{H_2} , i_{HD} , and i_{D_2} are the components of the current which caused the evolution of the molecules H_2 , HD , and D_2 , respectively. The rate of formation of hydrogen molecules, which is proportional to i_{H_2} , is given by

$$d(H_2)/dt = i_{H_2} = kP_H([H^+] + [D^+]) \cdot (K_1 C_{H_2O} + K'_1 C_{HOD})$$

since the rate at which isotopic ions exchange with light hydrogen ions resulting from H_2O and HOD dissociations is given by the product of the last two terms in brackets. Similarly we obtain

$$i_{D_2} = kP_D([H^+] + [D^+]) \cdot (K_2 C_{HOD} + K'_2 C_{D_2O})$$

$$i_{HD} = kP_H([H^+] + [D^+]) \cdot (K'_1 C_{HOD} + K_2 C_{D_2O})$$

$$+ kP_D([H^+] + [D^+])(K_1 C_{H_2O} + K'_1 C_{HOD})$$

$$i_{H_2} = kP_H([H^+] + [D^+])(K_1 C_{H_2O} + K'_1 C_{HOD})$$

Substituting into the relation

$$(2i_{H_2} + i_{HD})/(2i_{D_2} + i_{HD}) = \alpha(1 - x)/x$$

we obtain

$$\alpha = 2K'_1/K_2 = K_1/2K'_2 = K_1/K_2 = 5.4$$

This result has been derived without considering the adsorption of hydrogen molecules at the cathode and hence is valid only for very small current densities. For larger currents the factor $1/(1 + \omega i)$ must be introduced; this changes the value of α .

In practically pure light water we find

$$\alpha_{H_2O} = \frac{K_1}{K_2} \cdot \frac{1}{1 + \omega i} = 2.7$$

and in practically pure heavy water

$$\alpha_{D_2O} = \frac{K_1}{K_2} (1 + \omega i) = 50$$

In a 50 per cent mixture

$$\alpha_{HOD} = 7$$

The separation coefficient on mercury in mixtures containing mainly light water has been found to lie between 2.7 and 2.8 and is well known to increase considerably at very high concentrations of heavy water.

It should be emphasized that the evolution of hydrogen discussed in the present article is that obtained in polarographic studies with very small current densities, not more than 0.001 ampere per square centimeter, so that no visible evolution of bubbles takes place. The cathodic interface

is, therefore, probably not supersaturated with molecular hydrogen. When bubbles become visible, no strict rules may be expected to hold.

REFERENCES

- (1) BOWDEN, F. P.: *Trans. Faraday Soc.* **24**, 474 (1928).
- (2) BOWDEN, F. P., AND KENYON, H. F.: *Nature* **135**, 105 (1935).
- (3) BOWDEN, F. P., AND RIDEAL, E. K.: *Proc. Roy. Soc. (London)* **A120**, 59 (1928).
- (4) DRUCKER, C.: *Trans. Faraday Soc.* **30**, 1071 (1934).
- (5) ERDEY-GRÚZ, T., AND VOLMBER, M.: *Z. physik. Chem.* **A157**, 182 (1931).
- (6) FRUMKIN, A.: *Z. physik. Chem.* **A164**, 121 (1933).
- (7) HERASYMENKO, P.: *Rec. trav. chim.* **44**, 503 (1925).
- (8) HERASYMENKO, P., AND ŠLENDYK, I.: *Z. physik. Chem.* **A149**, 123 (1930).
- (9) HEYROVSKÝ, J.: *Rec. trav. chim.* **44**, 499 (1925); **46**, 582 (1927).
- (10) HEYROVSKÝ, J.: *Actualités scientifiques et industrielles*, No. 90. Hermann et Cie., Paris (1934).
- (11) HEYROVSKÝ, J.: *Collection Czechoslov. Chem. Commun.* **9**, 273 (1937).
- (12) HEYROVSKÝ, J.: *Collection Czechoslov. Chem. Commun.* **9**, 345 (1937).
- (13) HEYROVSKÝ, J., AND BEREZICKÝ, S.: *Collection Czechoslov. Chem. Commun.* **1**, 28 (1929).
- (14) NOVÁK, J.: *Collection Czechoslov. Chem. Commun.* **9**, 207 (1937).
- (15) SCHWARZENBACH, G., EPPRECHT, A., AND ERLENMAYER, H.: *Helv. Chim. Acta* **19**, 1292 (1936).
- (16) SMITS, A.: *Z. physik. Chem.* **A172**, 470 (1935).
- (17) TOMEŠ, J.: *Collection Czechoslov. Chem. Commun.* **9**, 150 (1937).
- (18) TOPLEY, B., AND EYRING, H.: *J. Chem. Phys.* **2**, 217 (1935).
- (19) VOPIČKA, E.: *Collection Czechoslov. Chem. Commun.* **8**, 349 (1936).

THE CHEMISTRY OF RETENE

DAVID E. ADELSON¹ AND MARSTON TAYLOR BOGERT

Department of Chemistry, Columbia University, New York, New York

Received September 8, 1938

CONTENTS

| | |
|------------------------------------------------------|-----|
| I. Introduction..... | 135 |
| II. Natural sources of retene..... | 136 |
| III. Production of retene from natural sources..... | 137 |
| IV. Physical constants and properties of retene..... | 138 |
| V. Proof of the structure of retene..... | 138 |
| A. Oxidative degradation..... | 138 |
| B. Synthesis..... | 145 |
| VI. Hydrogenation products of retene..... | 148 |
| A. Various hydrogenated retenes..... | 148 |
| B. Relation to fichtelite..... | 149 |
| VII. Derivatives of retene..... | 151 |
| A. Hydroxy derivatives..... | 151 |
| B. Carboxylic acids..... | 157 |
| C. Sulfonic acids..... | 157 |
| D. Nitro and amino derivatives..... | 158 |
| E. Halogenated derivatives..... | 159 |
| F. Retenequinone..... | 160 |
| G. 6-Acetylretene..... | 164 |
| H. Polynuclear compounds..... | 166 |
| VIII. Conclusion..... | 170 |
| A. Relation of retene to the terpenes..... | 170 |
| B. Comparison of retene and phenanthrene..... | 171 |
| IX. References..... | 172 |

I. INTRODUCTION

It is interesting to note the cycle through which the organic chemist has passed in a century of progress during which he has synthesized several hundred thousand new compounds and has unravelled, step by step, many of nature's secrets that had long been considered insolvable. At first the organic chemist was concerned chiefly with the isolation of naturally occurring materials, some simple and some complex in structure, and the establishment of their structures by synthetic and degradative methods. This task being well on the way to solution, he busied himself with the synthesis and the proof of the structure of several hundreds of thousands of new compounds. In recent years, however, the discovery of the pres-

¹ National Research Fellow in Chemistry, Columbia University, 1936-37. Present address: Shell Development Company, Emeryville, California.

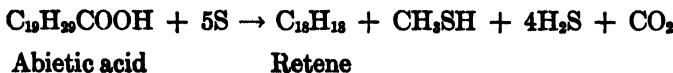
ence of the phenanthrene nucleus in the sterols, bile acids, cardiac poisons, sex hormones, alkaloids of the morphine and aporphine groups, many carcinogenic hydrocarbons and in a variety of other materials has prompted the organic chemist to investigate once again naturally occurring substances. As a result of this revival of interest in products of natural occurrence, much synthetic activity has been undertaken involving such substances as intermediates or as starting materials.

Because of the current interest in naturally occurring compounds, this review of the chemistry of retene has been written. Retene is itself of further interest on account of its structural relationship to the resin acids. The hydrocarbon is worthy of additional consideration by virtue of its possession of the biologically significant methyl and isopropyl groups in positions similar to their occurrence in many terpenes, camphor, and the *p*-cymenes.

II. NATURAL SOURCES OF RETENE

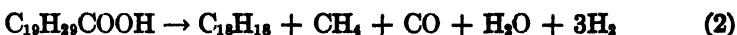
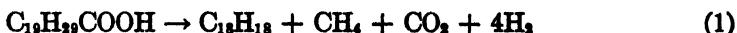
Retene (Greek, *rhētinē*, pine resin) was first described by Fikentscher and Trommsdorf (140) a century ago. Their source of the hydrocarbon was fossilized pine wood in a peat bed of the Fichtelgebirge region of Bavaria. In 1858 retene was further investigated by Fehling (55) and Fritzsche (67, 69), who studied a sample of the relatively pure hydrocarbon obtained by Knauss from the oil resulting from the distillation of pine wood. Wahlforss (146) and Ekstrand (51) secured retene by allowing the oil from the higher boiling fractions of various wood-tar oils to deposit the hydrocarbon and then pressing the cold oil. In 1887 processes were developed and patented for the production of retene by heating resin oils with sulfur until the evolution of hydrogen sulfide had ceased and recovering the hydrocarbon from the product by distillation or extraction with suitable solvents (7, 37, 94).

Vesterberg (143) obtained retene by treating abietic acid with sulfur and distilling the product at reduced pressure. Virtanen (145) secured the hydrocarbon in similar fashion from pinabietic acid. The work of Vesterberg and Virtanen served not only to introduce a method for the dehydrogenation of a hydroaromatic compound to an aromatic one, a procedure that has been adapted to many studies of natural products in recent years, but also to cast light on the structure of resin acids of the abietic type. Since the starting material contains twenty carbon atoms, the dehydrogenation to retene ($C_{18}H_{18}$) is accompanied by a loss of two carbon atoms and is probably that represented by the following equation:



An extensive study of the action of sulfur on rosin was made by Li Man Cheung (103), who encountered considerable difficulty in isolating in pure condition the retene thus formed.

Other dehydrogenating agents, such as selenium (46) and palladium charcoal (127), gave better yields of retene. Thus, while Vesterberg (143) using sulfur was able to get a yield of only 8 per cent, based on the abietic acid, the selenium treatment gave a yield of 55 per cent of the pure hydrocarbon. The palladium charcoal dehydrogenation, carried out by Ruzicka and Waldmann (127) at 300–330°C., gave a 71 per cent yield of fairly pure retene together with approximately 4 moles of hydrogen, 1 mole of methane, 0.75 mole of carbon dioxide, and 0.25 mole of carbon monoxide. These authors postulated the following dehydrogenation reactions as taking place simultaneously, reaction 1 taking place to a much greater extent than reaction 2:



III. PRODUCTION OF RETENE FROM NATURAL SOURCES

Two of the processes mentioned above have been adapted to large-scale production and have been patented. The essential difference between these two methods consists in the fact that one involves the use of sulfur as a dehydrogenation agent, while the second employs only a metallic halide as a catalyst in accomplishing the same end.

The sulfur dehydrogenation process (85) involves heating rosin, abietic acid, esters of abietic acid, abietine, abietene, or abietane with sulfur at about 250°C. for a short time and then distilling the reaction mass under reduced pressure (25 to 28 in. of mercury) and at an elevated temperature (250° to 275°C.) with steam. The oily distillate is dissolved in alcohol, and retene crystallizes out on cooling. A further quantity of retene may be obtained by evaporating the mother liquors, combining the residue with the highest boiling fractions of the original distillation, and subjecting the whole once more to the sulfur treatment.

The second process (113) comprises condensing pine-tar oils (the water-insoluble portion of the liquid obtained by destructive distillation of the wood) by heating them to about 250°C. with a metallic halide catalyst, of which aluminum chloride was found to be the best, and then distilling under a pressure of about 2 in. of mercury. The distillate is agitated with 10 per cent caustic soda solution, and the oily layer is separated and distilled at atmospheric pressure. The last 25 per cent of the distillate contains retene and other hydrocarbons. This is allowed to crystallize, and the retene is recovered by removing the admixed oil by hydraulic pressure of about 400 pounds per square inch.

IV. PHYSICAL CONSTANTS AND PROPERTIES OF RETENE

Retene crystallizes from alcohol in large, colorless plates, melting at 98.5–99°C. (133); it boils at 390°C. (29). It is characterized by its picrate ($C_{18}H_{18} + C_6H_5O_2N_3$), which forms orange-yellow needles melting at 123–124°C. (53, 68), and by its *sym*-trinitrobenzene adduct ($C_{18}H_{18} + C_6H_5O_2N_3$), which is obtained as yellow needles melting at 139–140°C. (139).

Retene forms a complex with 2-chloro-1,3,5-trinitrobenzene ($C_{18}H_{18} + C_6H_5O_2N_3Cl$), m. p. 53.5°C. (89). It forms two eutectic mixtures melting at 45.5°C. and 47.5°C. and containing 28.8 and 65.5 per cent of retene, respectively. Thermal analysis of the systems retene–2-chloro-1,3,5-trinitrobenzene, retene–2,4,6-trinitrophenol (88), retene–2,4,6-trinitroresorcinol (91), and retene–2,4,6-trinitroaniline (90) have been reported in the literature.

The hydrocarbon is somewhat volatile with steam (66, 67, 68). It sublimes in a vacuum of cathode light intensity at 135°C. (99); when maintained in such a vacuum for 14 hr. at 36°C., 2.2 per cent of the retene present will sublime (71). Its refractive power in benzene solution is given by Chilesotti (41). Its luminescence phenomena under the influence of cathode rays have been described by Pochettino (118). The fluorescence phenomena of retene under the influence of x-rays have been studied by Straus (138). When irradiated with ultraviolet light at the temperature of liquid air, retene emits a band spectrum (72). Wiedemann and Schmidt (152) observed in retene vapor a blue or violet fluorescence; in solution retene showed absorption in the extreme violet.

Ekstrand (51, 52, 53) gives the density of retene in crystalline form as 1.13, and after melting and supercooling as 1.08. In 100 parts of 95 per cent alcohol 69 parts of retene dissolve at the boiling point, while 3 parts dissolve at room temperature. The hydrocarbon is soluble in ether, carbon bisulfide, petroleum ether, benzene, and hot acetic acid.

The molar heat of combustion of retene at constant volume and at constant pressure has been measured by several investigators (30, 31, 137).

The behavior of retene in concentrated sulfuric acid in the presence of formaldehyde has been investigated by Ditz (49).

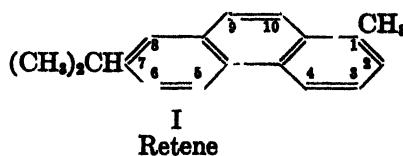
V. PROOF OF THE STRUCTURE OF RETENE

A. Oxidative degradation

Determination of the structure of retene proved to be a most perplexing problem, because of the great difficulty with which retene and its deriva-

tives burn when subjected to combustion in the ordinary way. This phenomenon, which was unnoticed for years, led to erroneous results which, in turn, furnished contradictory evidence as to the structure of the hydrocarbon. The combustion problem was solved by Bamberger and Hooker (19, 20, 21, 22), who burned a mixture of the material and lead chromate in a tube packed with lead chromate. Other investigators (5, 74) have found the analysis of retene derivatives to be both difficult and troublesome.

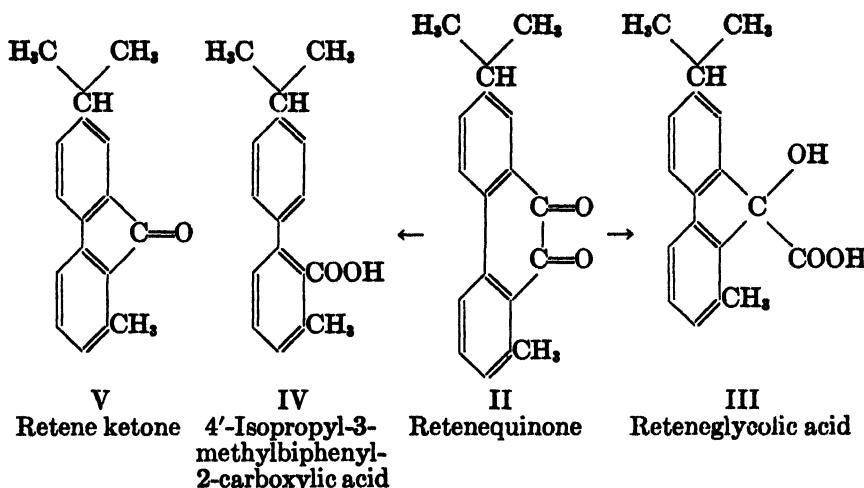
Retene (I) is 1-methyl-7-isopropylphenanthrene. It may be regarded as



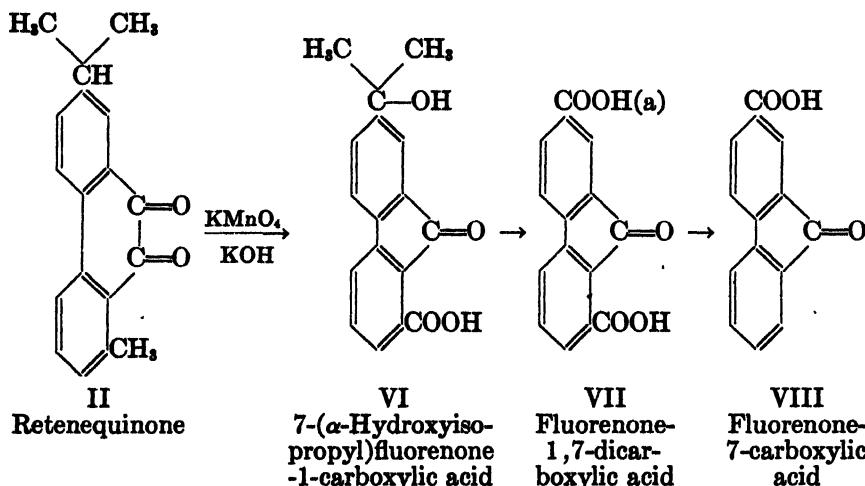
a methyl isopropyl homolog of biphenyl with an ethylene bridge inserted between the *o*- and *o'*-positions, one of these being adjacent to the carbon atom carrying the methyl group. Retene being a phenanthrene homolog, the arrangement of its double bonds is similar to that of phenanthrene, whose structure has been determined by Fittig and Ostermayer (62) and by Schultz (132).

Analyses indicate that retene possesses the empirical formula $C_{18}H_{18}$. Inasmuch as much of the early work on the proof of structure of the hydrocarbon led to unreliable conclusions because of erroneous analytical results and other difficulties, that work will not be included in this review.

After surmounting the combustion difficulties mentioned above, Bamberger and Hooker (15, 19) oxidized retene and obtained retenequinone (II), which had previously been prepared by Wahlforss (146). This was evidently an ortho-quinone, since Bamberger was able to form a quinoxaline derivative from it and to reduce it to a colorless hydroquinone. Distillation of the quinone with zinc dust yielded retene. Retenequinone (II) was further shown to be similar to an open chain α -diketone because of its sensitivity to alkali. When boiled with 10 per cent methyl alcoholic potassium hydroxide, it underwent the benzilic acid rearrangement to yield reteneglycolic acid (III) (104), and when fused with potassium hydroxide and lead peroxide, it gave 4'-isopropyl-3-methylbiphenyl-2-carboxylic acid (IV). The latter yielded retene ketone (V) when treated with thionyl chloride or when its barium salt was distilled with sodium methoxide.

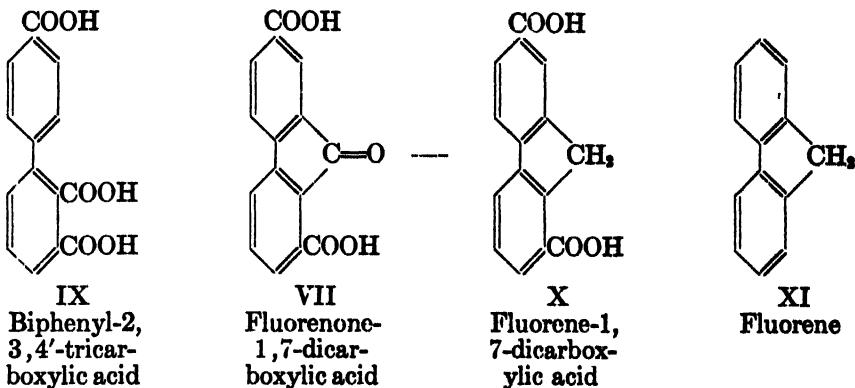


Bamberger and Hooker (19, 20, 21, 22) oxidized retenequinone (II) by means of alkaline potassium permanganate and secured a hydroxyisopropylidiphenyleneketonecarboxylic acid (VI), $C_{17}H_{14}O_4$. Further oxidation of VI by means of potassium dichromate in acid solution gave rise to a diphenyleneketonedicarboxylic acid (VII), which lost a molecule of carbon dioxide on distillation to give a diphenyleneketonecarboxylic acid (VIII).

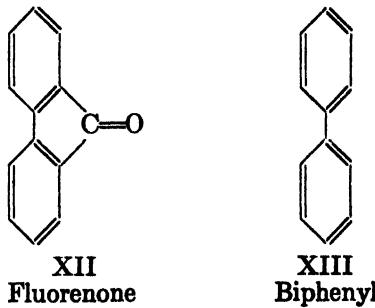


When the diphenyleneketonedicarboxylic acid (VII) was fused with potassium hydroxide, a diphenyltricarboxylic acid (IX) was formed, and when

reduced with sodium amalgam, a fluorenedicarboxylic acid (X) was obtained. Distillation of the latter with lime yielded fluorene (XI). The



dicarboxylic acid (VII) yielded fluorenone (XII) on complete decarboxylation, and the tricarboxylic acid (IX) gave biphenyl (XIII) on similar treatment.



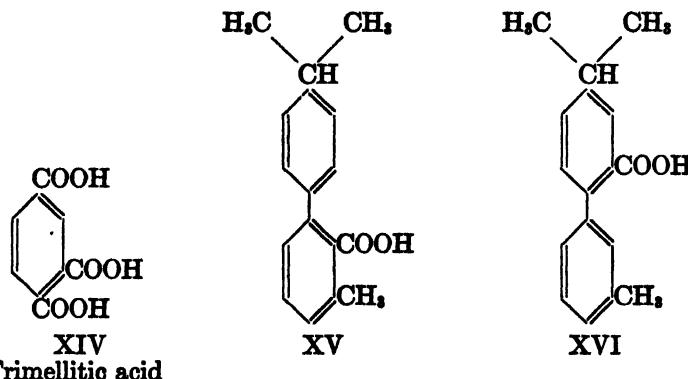
While the work of Bamberger and Hooker did not establish the exact structure of retene, it served to define certain important facts. The formation of a dicarboxylic acid (VII) showed that the C_4H_{10} -residue of retenequinone (II) must be in the form of two alkyl side chains. The possibilities are diethyl, methyl *n*-propyl, and methyl isopropyl. The production of the hydroxyisopropyldiphenyleneketonecarboxylic acid (VI) removed the first two possibilities. Since VI can be converted to fluorenone (XII), it is apparent that the oxidation of retenequinone, $C_{18}H_{16}O_2$ (II), to $C_{17}H_{14}O_4$ (VI) left the four carbon atoms of the two original alkyl groups intact. The carboxyl group of VI could not have come from an ethyl or propyl group without the loss of carbon, and it must therefore have arisen from the oxidation of a methyl group. The other alkyl group in retenequinone (II) is therefore *n*-propyl or isopropyl. The keto acid

(VI), however, contains one oxygen atom not accounted for in the keto and carboxyl groups. Since the *n*-propyl group is not susceptible to partial oxidation, whereas alkaline permanganate is known to hydroxylate an isopropyl group, the second alkyl group is an isopropyl group. The product of the oxidation of retenequinone (II) by means of alkaline potassium permanganate is therefore a hydroxyisopropyldiphenyleneketonecarboxylic acid (VI). The tricarboxylic acid (IX) readily forms an anhydride, while the dicarboxylic acid (VII) does not. One carboxyl group of the acid (IX) must be ortho to the carbonyl group of VII, and thus ortho to the carboxyl group which is produced in IX by fusion of VII with alkali. It follows that one alkyl group of retenequinone (II) is attached to C₁, adjacent to one of the quinoidal carboxyl groups.

Bamberger and Hooker therefore established retene as a methylisopropylphenanthrene in which one of the alkyl groups is at C₁. Their work did not indicate whether or not the second alkyl group was in the same benzene nucleus or in a different one.

Fortner (64) oxidized fluorene-2-carboxylic acid with chromic acid and obtained Bamberger and Hooker's diphenyleneketonecarboxylic acid (VIII). This showed further that one of the alkyl groups in retene was at C₇.

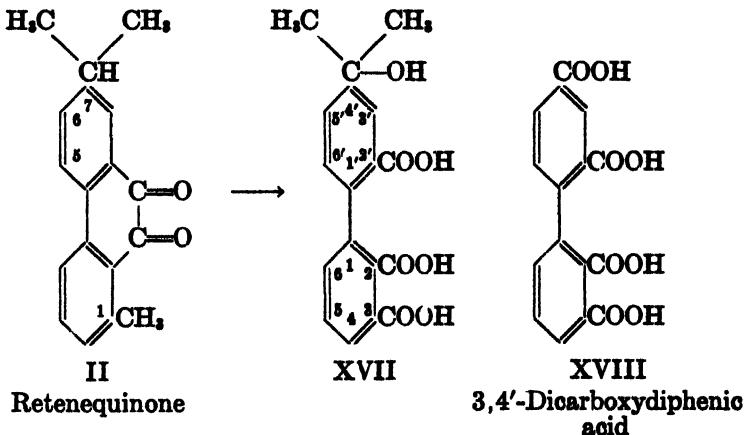
By fusion of retene ketone (V) with potassium hydroxide, Lux (104) obtained two methylisopropylbiphenylcarboxylic acids, XV and XVI. One of these acids, presumably XV, did not esterify easily, and he concluded that it must contain two substituting groups in the positions ortho to the carboxyl group. This lent further proof to the presence of one alkyl group in retene at C₁.



Schultze (133) oxidized hydrogenated retenes to trimellitic acid (XIV), thus proving that the two alkyl groups in retene cannot be in the same benzenoid ring. Since one of the alkyl groups in retene had already been shown to be at C₁, the work of Schultze indicated that the second alkyl

group was at C₇. Thus the observations of Fortner, Lux, and Schultze showed that the methyl and isopropyl groups in the retene molecule are probably located at C₁ and C₇, respectively, or at C₇ and C₁.

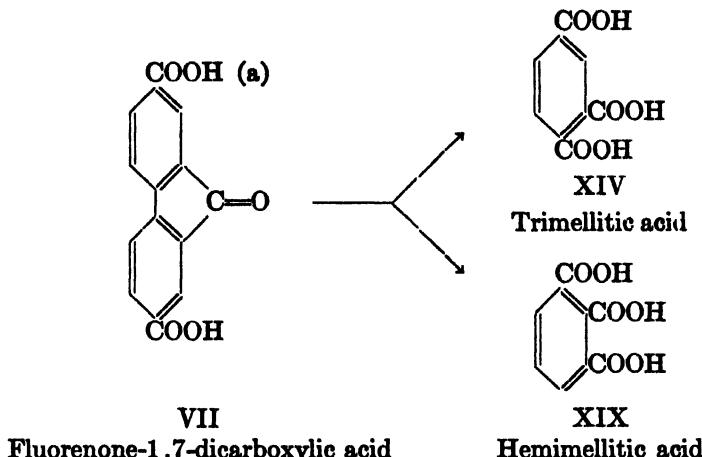
It remained for Bucher to complete the task of determining the structure of retene. Retenequinone (II) is oxidized by means of potassium permanganate in pyridine solution to yield 4'-(α -hydroxyisopropyl)-3-carboxy diphenic acid (XVII). This acid (XVII) loses water very easily upon heating, leaving a residue which is still soluble in sodium carbonate solution. This shows the acid to be tribasic. Its formation from retenequinone establishes the positions of two of the carboxyl groups at C₃ and C_{2'} (XVII). The loss of water on heating fixes another carboxyl group in an adjacent ortho position (C₂), for substances of the diphenic acid type sublime unchanged when heated. Inasmuch as the isopropyl group remains in the compound, it is certain that the carboxyl group attached to C₃ must have come from the methyl group present in the original quinone (II) at C₁.



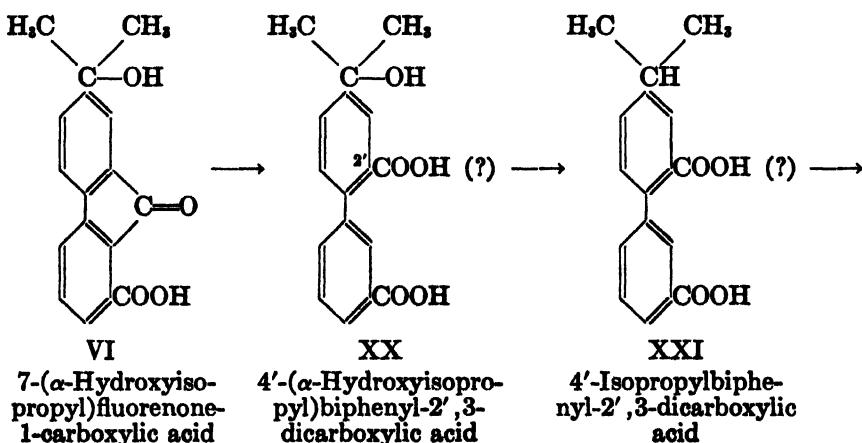
Further oxidation of the tribasic hydroxy acid (XVII) yields the corresponding tetrabasic acid (XVIII), 3,4'-dicarboxy diphenic acid, which gives biphenyl (XIII) on complete decarboxylation. This tetrabasic acid, on heating, gives off water to form an anhydride which is still soluble in sodium carbonate solution with effervescence of carbon dioxide. The acid (XVIII) therefore does not contain a carboxyl group at C_{2'}. In this oxidation a small quantity of hemimellitic acid (XIX) was obtained, proving that in retene the alkyl groups must be attached to different benzene rings. These results show that the isopropyl group in retenequinone (II) must be at C₁, C₆, or C₇.

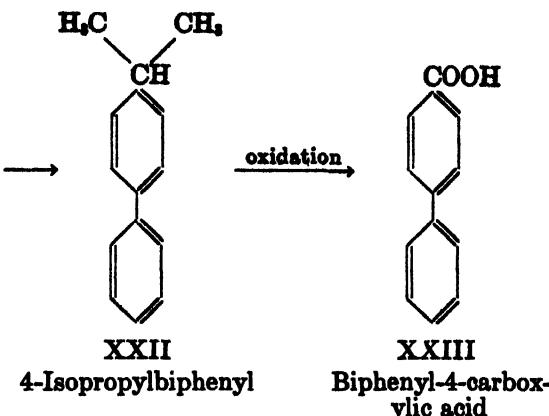
Additional evidence concerning the location of the isopropyl group was obtained by Bucher in his oxidation of the diphenyleneketonedicarboxylic

acid (VII) of Bamberger and Hooker to hemimellitic acid (XIX) and trimellitic acid (XIV). This further proves that the original alkyl groups in retene are in different benzene nuclei and that the unlocated carboxyl group (a) in the fluorenenedicarboxylic acid (VII), and hence the isopropyl group in retene, is at C₆ or C₇.

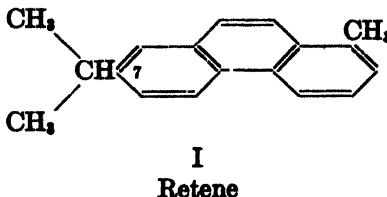


The isopropyl group in retene was finally established at C₇ by the degradation of the hydroxyisopropylfluorenonecarboxylic acid (VI) of Bamberger and Hooker to the known *p*-phenylbenzoic acid (biphenyl-4-carboxylic acid). Fusion of VI with potassium hydroxide gives the hydroxyisopropylbiphenyldicarboxylic acid (XX), which yields the isopropylbiphenyldicarboxylic acid (XXI) upon reduction with hydriodic acid. The position of the carboxyl group formed by opening the fluorenone ring of VI by alkali and represented at C₂, in formulas XX and XXI is





problematical; it may be at C₄. The exact position of this carboxyl group is of no serious consequence in the proof of structure in question, for this carboxyl group disappears, together with the other carboxyl group, upon decarboxylation of XXI. Oxidation of the 4-isopropylbiphenyl (XXII) thus formed gives *p*-phenylbenzoic acid (biphenyl-4-carboxylic acid) (XXIII), proving that the isopropyl group of retene is at C₇. Retene (I) is therefore 1-methyl-7-isopropylphenanthrene.

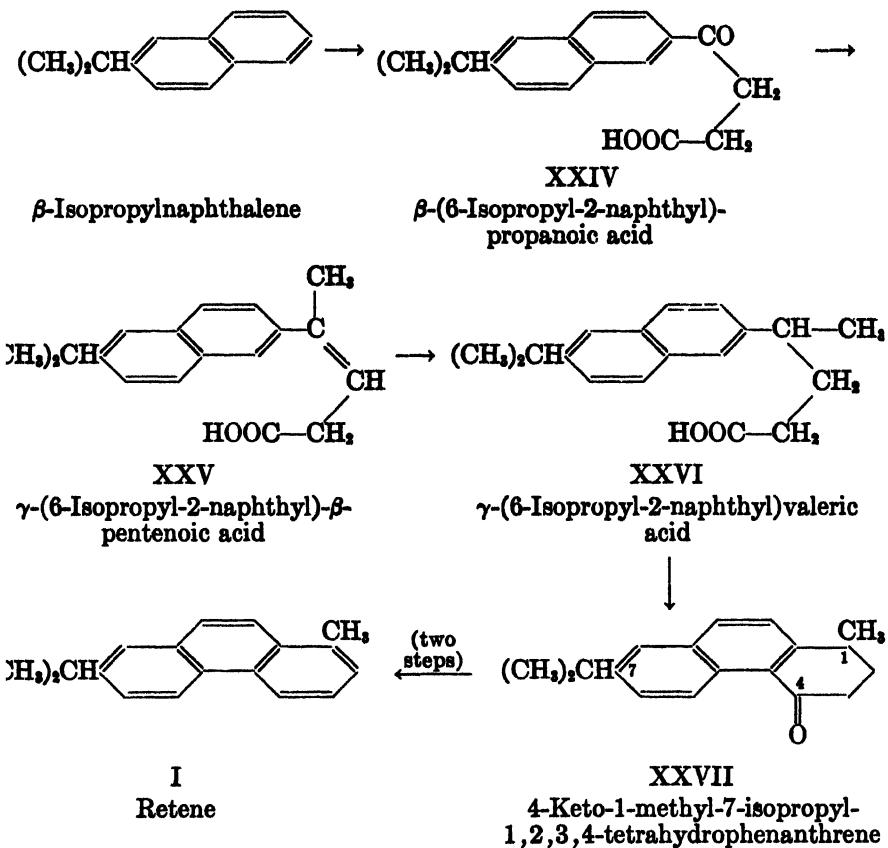


B. Synthesis

The structure of retene as established by Bucher has been further corroborated by the synthesis of the hydrocarbon from simpler materials of known constitution. Aside from their structural contribution, these syntheses have no practical value as a source of retene.

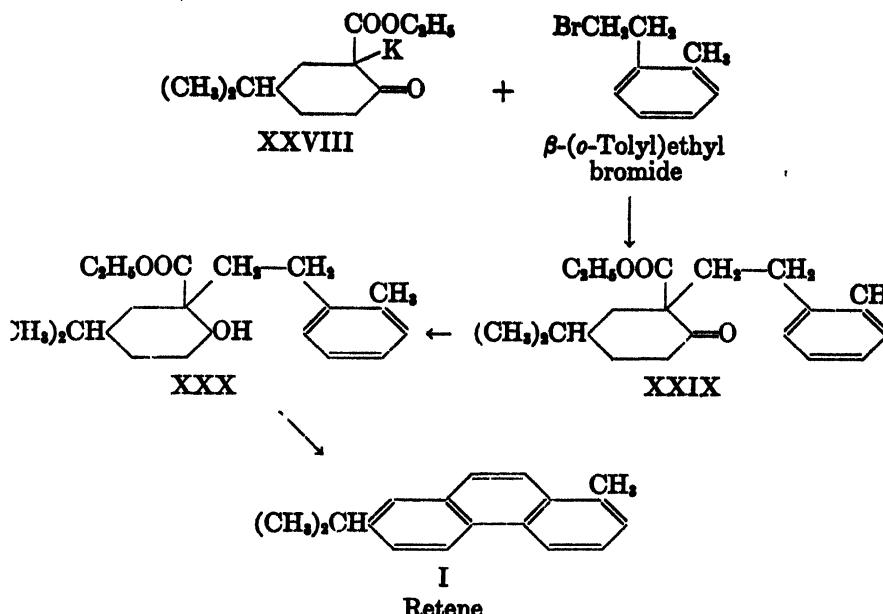
The succinic anhydride synthesis of Haworth (76) was adapted by that author, in collaboration with Letsky and Mavin (78), to the synthesis of retene (I). β -Isopropynaphthalene condenses with succinic anhydride in the presence of aluminum chloride to yield β -(6-isopropyl-2-naphthoyl)-propanoic acid (XXIV). Methylmagnesium iodide reacts selectively with the keto group of the methyl ester of the keto acid (XXIV) to give γ -(6-isopropyl-2-naphthyl)- β -pentenoic acid (XXV), which is reduced with phosphorus and hydriodic acid to γ -(6-isopropyl-2-naphthyl)valeric acid (XXVI). Cyclization of the latter by means of sulfuric acid yields

4-keto-1-methyl-7-isopropyl-1,2,3,4-tetrahydronaphthalene (XXVII). The Clemmensen reduction of the ketone followed by selenium dehydrogenation gives retene (I).

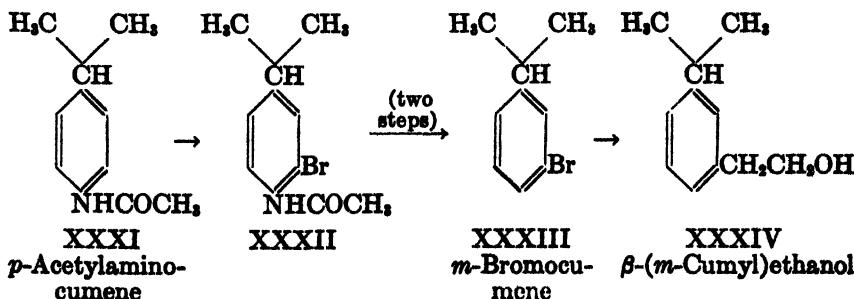


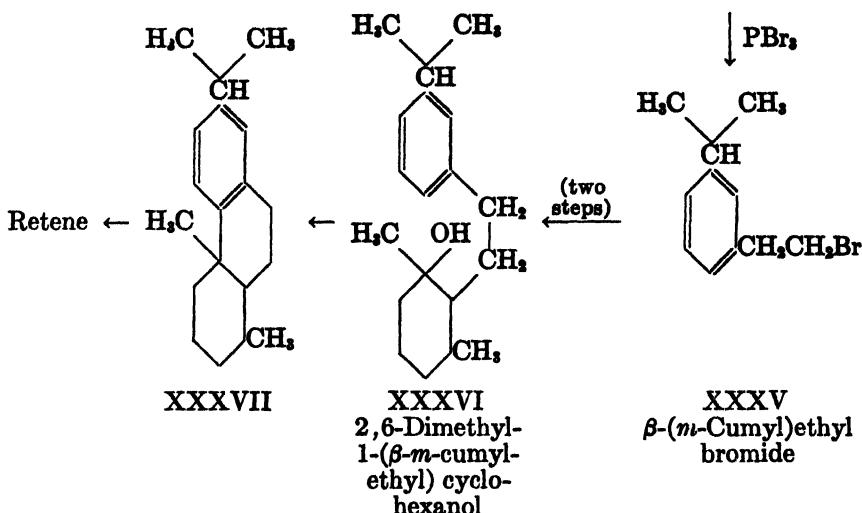
The Bardhan and Sengupta synthesis of phenanthrene (25) was adapted with equal success by the Indian chemists in the synthesis of retene (26). The potassium derivative of ethyl 4-isopropylcyclohexan-1-one-2-carboxylate (XXVIII) reacts with β -o-tolylethyl bromide to give ethyl 4-isopropyl-2-(β -o-tolylethyl)cyclohexan-1-one-2-carboxylate (XXIX). Sodium amalgam reduces the keto ester (XXIX) to ethyl 4-isopropyl-2-(β -o-tolylethyl)-cyclohexan-1-hydroxy-2-carboxylate (XXX). This hydroxy ester (XXX) is heated with phosphorus pentoxide and the product dehydrogenated by selenium to yield retene (I).

More recently Bogert and Sterling (35) have adapted the Perlman-Davidson-Bogert synthesis of phenanthrene (116) to the synthesis of retene. This method starts with the known *p*-acetylaminocumene



(XXXI), which is converted to the monobromo derivative. Hydrolysis of the latter, followed by removal of the amino group through the diazo reaction, gives *m*-bromocumene (XXXIII). The bromo compound reacts with magnesium, then with ethylene oxide to yield β -(*m*-cumyl)ethanol (XXXIV). The alcohol (XXXIV) gives the corresponding bromide (XXXV) when treated with phosphorus tribromide. The magnesium compound of this bromide undergoes the Grignard reaction with 2,6-dimethylcyclohexanone to form 2,6-dimethyl-1-(β -*m*-cumylethyl)cyclohexanol (XXXVI). Cyclization of the latter by means of sulfuric acid furnishes 12-methyl-1,2,3,4,9,10,11,12-octahydroretene (XXXVII), which is smoothly dehydrogenated by means of selenium to retene (I).





VI. HYDROGENATION PRODUCTS OF RETENE

A. Various hydrogenated retenes

The study of the hydrogenation products of retene has been of interest chiefly in connection with attempts to establish the structure of abietic acid. Retene having been defined as a dehydrogenation product of this acid, the ring skeleton of abietic acid is known. Many investigators (100, 133, 141) believed that, in order to determine whether hydrogenated retenes or their methyl homologs are obtained from resin acids, it was important to prepare the different hydrogenated retenes and determine their properties.

The earliest work on the hydrogenation of retene was done by Bamberger and Lodter (23), who reduced retene by means of sodium and amyl alcohol and obtained a tetrahydroretene. By prolonged action of phosphorus and hydriodic acid on the hydrocarbon, Liebermann and Spiegel (101) obtained a dodecahydroretene, $\text{C}_{18}\text{H}_{30}$. This compound was later prepared by Ipatiev (87), who hydrogenated retene at an elevated temperature in the presence of nickel oxide. Prolonged hydrogenation of this dodecahydroretene yielded perhydroretene, $\text{C}_{18}\text{H}_{32}$. The latter did not prove to be identical with fichtelite, in disagreement with the views of Spiegel (136) and of Bamberger and Strasser (24). In 1904 Easterfield and Bagley (50) heated abietic acid with hydriodic acid in a sealed tube and obtained a hydrocarbon which they reported to be a decahydroretene, $\text{C}_{18}\text{H}_{28}$.

In 1920 Virtanen (144) prepared six hydrogenated compounds from

retene and recorded their physical constants. Dihydroretene and tetrahydroretene (table 1) were made by reducing retene with sodium and amyl alcohol; the others were obtained by the action of phosphorus and hydriodic acid on the hydrocarbon at elevated temperatures in a sealed tube filled with carbon dioxide. None of these hydrogenated retenes formed picrates.

Czerny (45) obtained an octahydroretene, $C_{18}H_{36}$, by prolonged treatment of pine oil with caustic potash at 50–60°C.; its physical constants differed somewhat from those of the octahydroretene described by Virtanen (144). Orlov (115) investigated the pyrolytic dissociation of retene in the presence of hydrogen under pressure. The hydrogenation of retene and hydroretenes has been studied by Hasselstrom and Hull (75).

In 1932 Bogert and Hasselstrom (33) reduced retene-6-carboxylic acid, the synthesis of which will be described in a later section, and obtained an

TABLE I
Physical properties of retene and of its hydrogenation products

| COMPOUND | FORMULA | APPEARANCE | MELT- | BOILING | d_4^{20} | n_D^{∞} |
|----------------------|----------------|-----------------|--------------|--------------------|------------|----------------|
| | | | ING POINT | POINT AT 10 MM. | | |
| Retene..... | $C_{18}H_{16}$ | White plates | 98.5 | 208–210 | | |
| Dihydroretene..... | $C_{18}H_{20}$ | White plates | 64–65 | 188–190 | | |
| Tetrahydroretene.... | $C_{18}H_{22}$ | Pale yellow oil | | 180–183 | 1.0057 | 1.56061 |
| Hexahydroretene.... | $C_{18}H_{24}$ | Colorless oil | | 175–177 | 0.9802 | 1.54705 |
| Octahydroretene.... | $C_{18}H_{26}$ | Colorless oil | | 163–165 | 0.9578 | 1.53023 |
| Decahydroretene.... | $C_{18}H_{28}$ | Colorless oil | | 155–158 | 0.9342 | 1.51501 |
| Dodecahydroretene.. | $C_{18}H_{30}$ | Colorless oil | | 148–150 | 0.8985 | 1.48510 |

octahydroretene-6-carboxylic acid. Its sodium salt behaved as a true resinate, being soluble in hot water but not in aqueous alkali. This octahydroretene-6-carboxylic acid is the nearest approach so far achieved synthetically to a true aporesin acid of the abietic series.

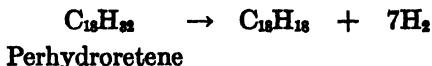
Owing to the difficulty encountered in the preparation of retene derivatives, Nyman (114) studied substitution in 9,10-dihydroretene and found that it occurred more readily than in retene. The dihydro compound was prepared in 60 per cent yield by reduction of retene with sodium and amyl alcohol. Nyman investigated the acetylation of 9,10-dihydroretene and prepared a number of new derivatives.

B. Relation to fichtelite

Fichtelite (XXXVIII), m. p. 46.5°C., is a completely saturated hydrocarbon which is decidedly inert in both chemical and physical properties. First isolated by Bromies (38) from peat beds of pine forests in Bavaria,

its structure and composition were long in question. Early investigators (17, 24, 84, 101, 136) regarded it as a perhydroretene, $C_{18}H_{32}$.

Prompted by the discovery that fichtelite yields retene upon dehydrogenation with sulfur (120), Ruzicka and Waldmann (128) studied a quantitative dehydrogenation of the material with palladium charcoal at 330–370°C. Were fichtelite a perhydroretene, $C_{18}H_{32}$, the gas evolved in this dehydrogenation should consist only of hydrogen:



The gases actually evolved were found to consist of one volume of methane and six volumes of hydrogen:

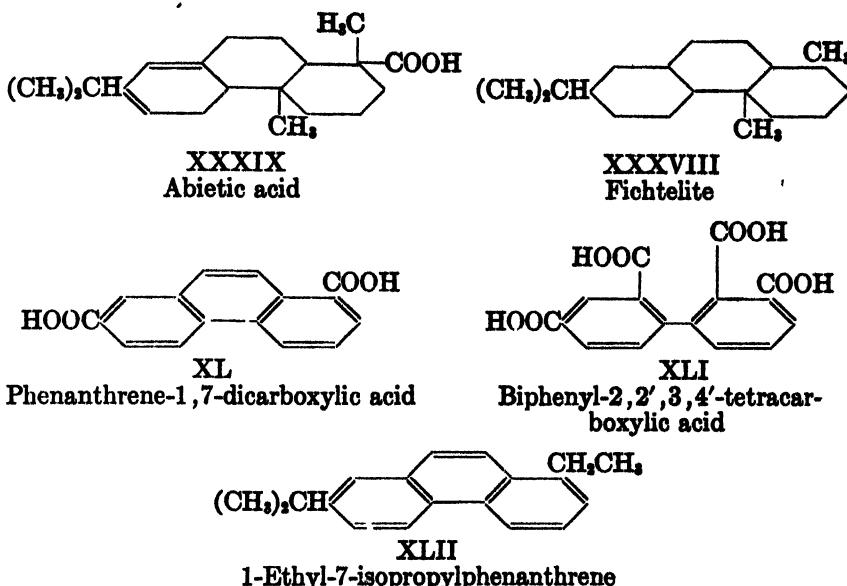


This indicated the presence of an additional tertiary methyl group. In order to show its relationship to abietic acid (XXXIX), from which it probably arises in the process of decay, Ruzicka and Waldmann have proposed for fichtelite the structure given in formula XXXVIII.

In an effort to clear up the structure of fichtelite, Bogert and Sterling (35) synthesized 12-methyl-1,2,3,4,9,10,11,12-octahydroretene (XXXVII), which was catalytically hydrogenated at 225°C. and 150 atm. pressure by Dr. Homer Adkins. There was obtained a hydrocarbon, $C_{19}H_{34}$, which was an oil, boiling at 179–181°C. at 12 mm. pressure. As fichtelite is a white crystalline solid, m. p. 46.5°C., this hydrocarbon was not identical therewith. It may be that the difference between the two is a stereochemical one (120) or that the hydrocarbon of Bogert and Sterling requires further purification.

It is of interest to point out at this point that the so-called "methylretene" obtained by Ruzicka and Meyer (124) in their degradation of ethyl abietate to "methylretene" has been identified as 1-ethyl-7-isopropylphenanthrene (XLII). Upon oxidation with alkaline potassium ferricyanide, retene yielded phenanthrene-1,7-dicarboxylic acid (XL) and biphenyl-2,2',3,4'-tetracarboxylic acid (XLI) (121). A "methylretene" should yield a tribasic acid under similar conditions.

Instead, Ruzicka, de Graaf, and Müller obtained phenanthrene-1,7-dicarboxylic acid (XL), identical with the product obtained from retene. This evidence indicates that the additional methyl group in "methylretene" is not attached directly to the phenanthrene ring. Ruzicka, de Graaf, and Müller postulated that the methyl group in question was in a side chain, probably united with the methyl group of retene (I) at C_1 in the form of an ethyl group. In other words, "methylretene" is 1-ethyl-7-isopropylphenanthrene (XLII).



This hypothesis was established as a fact by Haworth (77), who synthesized the hydrocarbon by a series of reactions similar to those used in his synthesis of retene (I) (see above). In the transformation from the methyl ester of the keto acid (XXIV to XXV), he used ethylmagnesium iodide instead of the methyl compound. The synthetic 1-ethyl-7-isopropyl-phenanthrene (XLII) was found to be identical with the "methylretene" obtained from the degradation of ethyl abietate (124).

VII. DERIVATIVES OF RETENE

Nearly one hundred years had elapsed since the preliminary report of Fikentscher and Trommsdorf (140) before systematic synthesis of retene derivatives was undertaken. Prior to the publication of Fieser and Young (61), no retene derivatives had been reported in which the exact position of all the substituent groups in the molecule was given. The unique work of these authors in establishing the structures of the two retenols of Komppa and Wahlforss (97) has largely made possible recent expansion in the retene field by other investigators. The matchless contribution of Fieser and Young (61) is discussed in detail below.

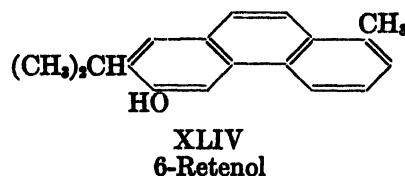
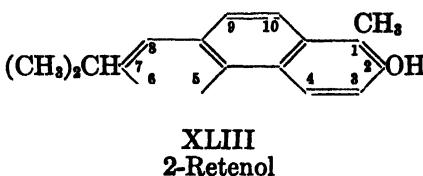
A. Hydroxy derivatives

By alkaline fusion of the ammonium salts of A- and B-retenesulfonic acids, Komppa and Wahlforss (97) prepared the corresponding mono-hydroxy derivatives. The term "retenol" was suggested for mono-hydroxyretene, $C_9H_{17}OH$, in analogy with the corresponding term for

monohydroxyphenanthrene. The A- and B-retenols of Komppa and Wahlforss (97) were characterized as 2- and 6-retenol, respectively, by the research of Fieser and Young (61).

Inasmuch as no promising method of degradation could be found to establish the structure of the two retenols, Fieser and Young tackled the problem from an entirely different angle. Since it is possible to oxidize suitable A- and B-retenol derivatives to 9,10-retenequinones without affecting the new substituents, these groups must occupy the 2-, 3-, 4-, 5-, 6-, or 8-positions. The 8-position is ruled out by the Dimroth test (47, 48) for α -hydroxyquinones. 8-Hydroxyretenequinone should give an intense coloration with boroacetic anhydride, but neither the A- nor the B-quinone displays such behavior. The 8-position is definitely ruled out by subsequent work of Adelson and Bogert (3, 4), who synthesized 3'-methyl-5,6-cyclopentenoretene from 6-acetylretene. The structure of 6-acetylretene had previously been established by these authors (1) by conversion of the ketone into the 6-retenol of Fieser and Young. Had the 8-position been the one involved, the cyclopenteno ring could not have been formed. The synthesis of 5,6-benzoretene by Adelson and Bogert (5) furnishes another example in substantiation of this point.

The next significant fact is that B-retenol couples with diazotized amines, while A-retenol does not. The double bonds of phenanthrene occupy fixed positions as in the case of naphthalene. C₂ and C₃ are connected by a single bond and hence a 2-phenanthrol cannot couple with diazonium salts if C₁ is blocked, as is the case with retene, for a double bond is required for the formation of the intermediate addition product. Stated in more general terms, this rule asserts that only those phenanthrols can couple which have a free position para to the hydroxyl group or have free an ortho position which is connected by a double bond to the carbon carrying the hydroxyl group. In the case of retene, there is only one possible hydroxyl derivative which does not fulfil the requirement. Since A-retenol does not couple, it must be 2-retenol (XLIII).



Four possible locations remain for the hydroxyl group in B-retenol, namely, 3, 4, 5, or 6. By a comparison of the reduction potentials of the 2- and B-hydroxyretenequinones, positions 4 and 5 are eliminated. Fieser (57) has shown that there is considerable difference in the effect of the

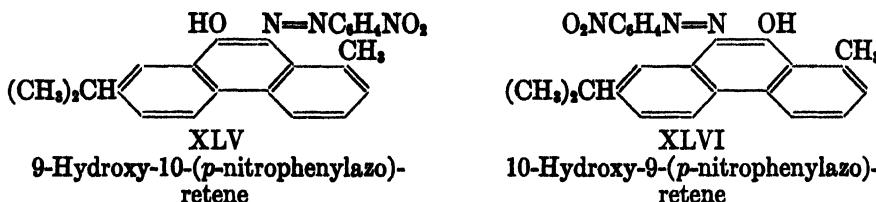
hydroxyl group on the potential of phenanthraquinone according to whether the substituent occupies a position which is meta or ortho-para to one of the ketonic oxygen atoms. A single hydroxyl group at C₂ or C₄ is without influence, while such a group at C₁ or C₃ lowers the potential 51 millivolts. In the case of the hydroxyretenequinones the situation is slightly altered by the fact that there must be present in one of the terminal benzene rings two substituents (hydroxyl and alkyl) which produce a lowering in the potential of the parent quinone. Comparing the mono- and di-hydroxyphenanthraquinones it is seen that the effect of a second hydroxyl group introduced into the ring carrying the first such group produces a somewhat greater potential lowering: 28 millivolts for a group in the meta position and -77 millivolts for a group in the ortho-para position. It is reasonable to expect that a hydroxyl group would have a similar effect upon the potential of retenequinone. The potentials of the three compounds concerned are:

| | POTENTIALS | Δ |
|--------------------------------|------------|------------|
| | volts | millivolts |
| Retenequinone | 0.421 | |
| A(2)-hydroxyretenequinone..... | 0.385 | -36 |
| B-hydroxyretenequinone..... | 0.348 | -73 |

The differences, which show the effect of the added hydroxyl groups, clearly indicate that the class A compounds contain the hydroxyl group in the meta position to one of the ketonic oxygen atoms in the appropriate quinone, and the class B compounds contain it in the ortho-para position. This confirms the structure already assigned to the A-compound (2-retenol) and shows that in B-retenol the hydroxyl group must be located at either C₃ or C₆ (C₈ is excluded by the Dimroth test). A choice between C₃ and C₆ can be made by a consideration of orientation principles. The retenols, it will be recalled, were made by alkali fusion of the corresponding sulfonates. It is highly improbable that an alkyl group (at C₁ or C₇) would direct a sulfonic acid group to a meta position (C₃ or C₅) when a position ortho to such a group is available. Since C₃ and C₄ are ruled out by other considerations, it is very probable that the hydroxyl group in B-retenol is at C₆ (XLIV).

9-Retenol has been made by the reduction of retenequinone with zinc dust and acetic acid (61). Its structure is established by a study of the 9,10(or 10,9)-hydroxyazo derivatives of retene. In the case of phenanthrene such a compound may be prepared in two ways: by coupling 9-phenanthrol with a diazotized amine or by condensing phenanthraquinone with phenylhydrazine. Werner and Frey (149) have shown that the same

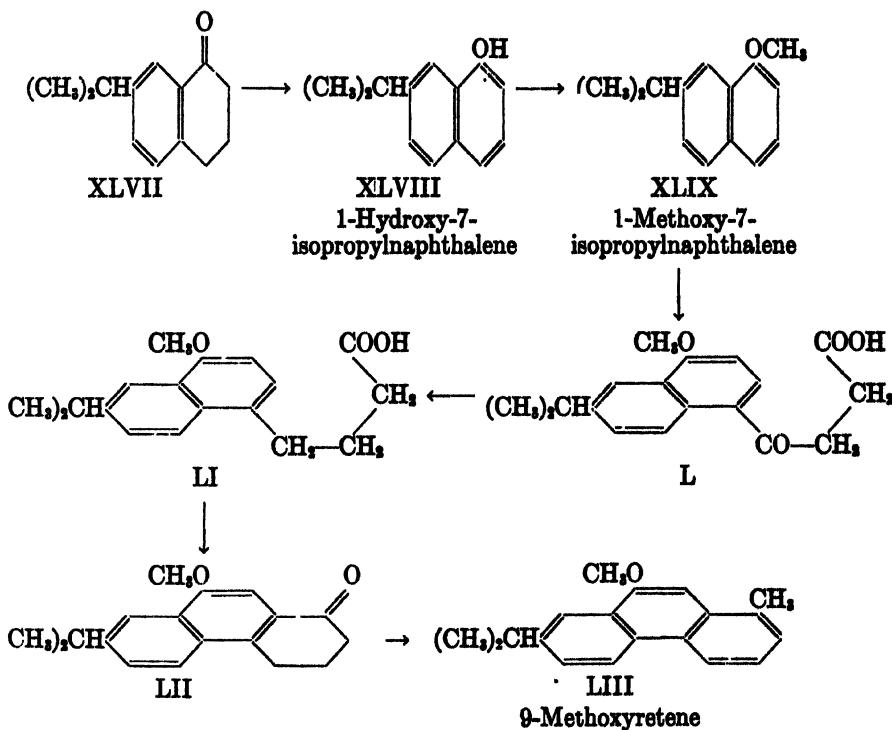
product is obtained in each case. With retene, however, the two reaction products are different. 9-Retanol couples readily with diazotized *p*-nitroaniline to give 9-hydroxy-10-(*p*-nitrophenylazo)retene (XLV). The isomeric compound, 10-hydroxy-9-(*p*-nitrophenylazo)retene (XLVI) has been prepared by Li Man Cheung (103) from retenequinone and *p*-nitrophenylhydrazine. Since this last reaction is probably one involving an



addition, and since it is known to be subject to steric influence, it is reasonable to assume that condensation at the ketonic group closest to the methyl group is retarded and that the product formed has the structure given in formula XLVI. The compound obtained from the retenol, being isomeric with this substance, must have the alternate structure of XLV, whence the new hydroxyretene is 9-retanol. Fieser and Young (61) assume that it is the methyl group at C₁ in retenequinone that causes the condensing agent (here *p*-nitrophenylhydrazine) to attack the ketonic group at C₉ rather than the one at C₁₀. Analogous experiments with 1-methylphenanthraquinone support this assumption.

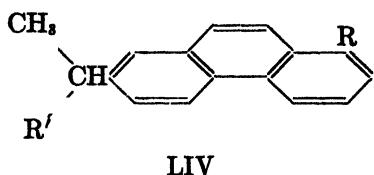
The remarkable reasoning of Fieser and Young (61) that it is the oxygen atom closest to the methyl group in retenequinone that is eliminated on reduction has recently been confirmed by synthesis. Keimatsu, Ishiguro, and Sumi (93) have synthesized 9-methoxyretene (LIII) by a process similar to that used by Ruzicka and Waldmann (126) in the preparation of 1,7-dimethyl-9-methoxyphenanthrene. Isopropylbenzene and succinic anhydride react in the presence of aluminum chloride in tetrachloroethane solution to give β -(*p*-isopropylbenzoyl)propanoic acid, which undergoes the Clemmensen reduction to yield γ -(*p*-isopropylphenyl)butanoic acid. The latter is converted into 1-keto-7-isopropyl-1,2,3,4-tetrahydronaphthalene (XLVII) by treatment with sulfuric acid. This keto compound (XLVII) furnishes 1-hydroxy-7-isopropynaphthalene (XLVIII) when heated with sulfur and cupric sulfide. 1-Methoxy-7-isopropynaphthalene (XLIX), prepared by the action of dimethyl sulfate on the hydroxy compound (XLVIII), yields β -(1-methoxy-7-isopropyl-naphthoyl-(4))-propanoic acid (L) when subjected to the Friedel-Crafts reaction with succinic anhydride. This keto acid (L) undergoes the

Clemmensen reduction to γ -(1-methoxy-7-isopropylnaphthyl-(4))-butanoic acid (LI), the acid chloride of which furnishes 1-keto-7-isopropyl-9-methoxy-1,2,3,4-tetrahydrophenanthrene (LII) upon vacuum distillation. The ketone (LII) is reacted with methylmagnesium iodide and the reaction product is dehydrogenated with selenium to yield 9-methoxyretene (LIII), identical with that prepared by Fieser and Young (61) by methylation of 9-retenol.

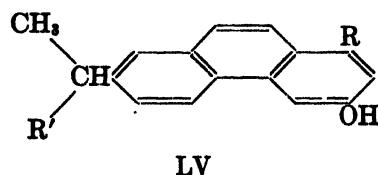


In 1933 Japanese chemists published the first of a series of papers concerning the isolation of a natural product which they subsequently described as possessing the retene nucleus. From the ether extract of the middle, red part of Japanese hinoki wood, Yoshiki and Ishiguro (154) obtained an oil which separated into two layers on standing. The upper layer, a viscous, bright yellow oil, yielded a crystalline substance which they called hinokiol. Subsequent work by Keimatsu and Ishiguro (92) gave $C_{14}H_{22}O_2$ as the molecular formula for hinokiol. Dehydrogenation of hinokiol with selenium gave three compounds which the authors state

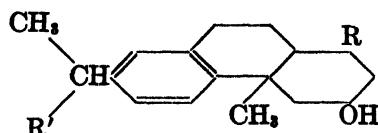
to be retene, a hydroxyretene, $C_{18}H_{18}O$, and a dihydroxyretene, $C_{18}H_{18}O_2$. Various derivatives were prepared in support of these views by Keimatsu and Ishiguro, who then proposed tentative structures for the compounds in question. The hydroxyretene, the dihydroxyretene, and hinokiol were assigned the structures LIV, LV, and LVI, respectively, where R is CH_3 and R' is CH_2OH , or R is CH_3OH and R' is CH_3 . The position of the phenolic hydroxyl group in LV and LVI was also uncertain; in addition, the double bond linkage and the position of the angular methyl group in hinokiol (LVI), which is a dihydroxyoctahydromethylretene, were not definite. Until a more comprehensive study of this problem appears, these structures are to be regarded as provisional.



LIV



LV

LVI
Hinokiol

9,10-Dihydroxyretene has been prepared by Bamberger (16) and by Knesch (96); the latter prepared it by reduction of the sodium bisulfite compound of retenequinone with iron or zinc dust and water, no acid or alkali being required. This hydroquinone is extremely sensitive to oxidation and is best isolated in the form of its diacetyl or dibenzoyl derivatives. Its instability had also been noticed by Ekstrand (53).

B. Carboxylic acids

Liebermann and Zsuffa (102) were the first to describe a retenecarboxylic acid. They synthesized it by the usual Friedel-Crafts reaction, from retene and oxalyl chloride, and reported its melting point as 121–123°C. It is possible that this melting point is a misprint for 221–223°C. This work was repeated by Komppa and Wahlfors (98), who obtained an acid, m. p. 229–231°C. (uncorrected) and oxidized it to the corresponding quinonecarboxylic acid, thus proving that the carboxyl group was not attached to C₉ or C₁₀. Bogert and Hasselstrom (33) obtained this retenecarboxylic acid, m. p. 237.5–238.5°C. (corrected), by the oxidation of acetylretene by sodium hypobromite.

The structure of this retenecarboxylic acid was established by Adelson and Bogert (1), who determined the position of the acetyl group in acetyl-retene as C₆ (see below). These authors (1) also simplified the synthesis of retene-6-carboxylic acid from 6-acetylretene by the use of an alkaline iodine-potassium iodide solution as the oxidant.

C. Sulfonic acids

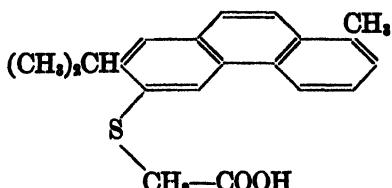
The action of sulfuric acid on retene has been studied by a number of investigators. Fehling (55) noted that sulfur dioxide was evolved when retene was heated with concentrated sulfuric acid above 100°C., but could isolate no sulfonic acid. Fritzsche (70) was successful in isolating the barium salt of a retenedisulfonic acid, C₁₈H₁₆(SO₃H)₂, obtained by the action of a mixture of concentrated and fuming sulfuric acids on retene. Ekstrand (53) was able to identify the free disulfonic acid and its chloride. All his attempts to prepare a monosulfonic acid, in spite of the use of the most diversified experimental conditions, were fruitless.

In 1930 Komppa and Wahlforss (98) prepared two isomeric retenesulfonic acids, C₁₈H₁₇SO₃H, designated as A-retenesulfonic acid and B-retenesulfonic acid, by adding concentrated sulfuric acid to vigorously stirred melted retene. At 100°C. the formation of the A-acid was favored, while at 200°C. the B-acid predominated. Fusion of the ammonium salts of these acids with potassium hydroxide gave the A- and B-retenols, whose structures have been established as 2- and 6-retenol, respectively, by Fieser and Young (61) (see above). Hence the acids in question are retene-2-sulfonic acid and retene-6-sulfonic acid.

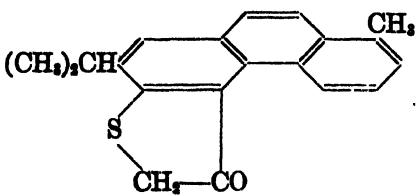
The work of Komppa and Wahlforss (98) on the retenesulfonic acids has been extended by Komppa and Fogelberg (97) to 9,10-dihydroretene. Sulfonation of 9,10-dihydroretene yields two well-defined sulfonic acids whose ethyl esters give upon oxidation the ethyl 2-retenequinonesulfonate and ethyl 6-retenequinonesulfonate of Komppa and Wahlforss (98). This defines the dihydroretenesulfonic acids as 9,10-dihydroretene-2-sulfonic acid and 9,10-dihydroretene-6-sulfonic acid. Attempts to replace the sulfonic acid group of these acids by the hydroxyl group results in simultaneous dehydrogenation and consequent formation of the corresponding retenols, the 2- and 6-retenols.

Hasselstrom and Bogert (74) have been interested in opening a path to a successful synthesis of thioindigoids from retene. 6-Retenesulfonyl chloride (98) was reduced by zinc and sulfuric acid to 6-thioretенол (6-retyl mercaptan), and this mercaptan was condensed with chloroacetic acid to 6-retylthioglycolic acid (I.VII). When this thioglycolic acid was subjected to the action of chlorosulfonic acid, there resulted a very un-

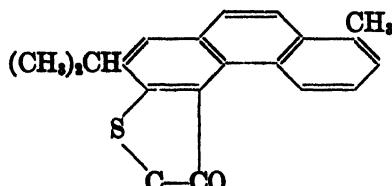
stable thioindoxyl (LVIII) and a disulfonic acid of the corresponding thioindigo. In alkaline solution the thioindoxyl (LVIII) was oxidized to an amorphous dark solid, presumably 6-retenethioindigo (LIX), which resisted purification and which, therefore, could not be positively identified. The retenethioindoxyl (LVIII) condensed readily at C_{1'} (formula LVIII) with various aldehydes to yield thioindogenides and with isatin to yield 6-retothiophene-3-indoleindigo. These condensation products, which also could not be purified, exhibited tinctorial properties.



LVII
6-Retylthioglycolic acid



LVIII



LIX
6-Retenethioindigo (?)

D. Nitro and amino derivatives

Attempts by various investigators to obtain crystalline nitro derivatives by the direct nitration of retene itself have not been very successful, the results in most cases having been either tars or oxidation products (22, 55, 68, 103). Ekstrand (53) obtained a yellow, flocculent precipitate when he diluted the acid mixture upon completion of the nitration of retene, but it resinsified when crystallization was attempted. Arnot (9) obtained a similar product, which he stated to be a dinitroretene, for he reduced it to a base (10) from which azo dyes were prepared by the usual reactions. He claimed also the production of a nitroretenesulfonic acid by nitration of retene dissolved in a 20 per cent oleum, and of a dinitroretenequinone from the quinone, but none of these products appears to have been either crystalline or analytically pure. A dinitroretenequinone has been reported by Li Man Cheung (103).

Heiduschka and Scheller (83) have reported a nitrotribromoretene-quinone, and Bogert and Hasselstrom (32) a crystalline dinitro-6-acetyl-retene. The latter authors (34) have also prepared mononitro derivatives of 6-acetoxyretene and of retene ketone (1-methyl-7-isopropylfluorenone), both of which were reduced to the corresponding amines. The amino-6-acetoxyretene was very unstable, but yielded a stable, crystalline acetyl derivative. Fieser and Young (61) found that 5-amino-6-retenol was similarly unstable, and they also resorted to acetylation to obtain a crystalline pure product. Bamberger and Hooker (22, page 144) have shown that 1-methyl-7-isopropylfluorene (retenefluorene) yields without difficulty a dinitro derivative. More recently Komppa and Wahlforss (98) have reported unsuccessful attempts to prepare a crystalline nitro derivative of retene.

From the foregoing it is apparent that, while retene itself yields no good nitration products, substituted retenes can be nitrated with success.

Only one amino derivative of retene is known, 6-aminoretene (LXXXIX) (1), a stable crystalline substance which is made by hydrolysis of 6-acetylaminoretene (LXXVIII), a compound resulting from the Beckmann rearrangement of 6-acetylretene oxime (LXXVII) (sec below).

E. Halogenated derivatives

The recorded literature contains but little systematic work on the halogen derivatives of retene. Such work was initiated by Ekstrand (53) in 1877. Ekstrand discussed the formation of additive compounds of retene with chlorine and bromine, and stated that the dichloride produced with chlorine appeared to be convertible into a monochlororetene. A dibromoretene was prepared by the action of bromine on retene in aqueous suspension, while a tetrabromo derivative was produced when excess bromine was used hot. Heiduschka and Scheller (83) oxidized tetrabromoretene and obtained a tribromoretenequinone, thus indicating that one of the bromine atoms in the original compound was at C₉ or C₁₀. These authors prepared a number of derivatives of this quinone.

By passing chlorine into a carbon tetrachloride solution of retene in the sunlight and in the presence of iodine, Heiduschka and Grimm (80) obtained a retenenonachloride, C₁₄H₆Cl₉. This was a white, amorphous powder, m. p. 98–100°C., which was soluble in the usual organic solvents and which was stable toward chromic acid and nitric acid in acetic acid solution.

Chlorination of retene in carbon tetrachloride solution gave 9(or 10)-chlororetene (98), which yielded retenequinone upon oxidation. Boiling the chlorinated product with alcoholic potassium hydroxide or with quinoline did not split off hydrogen chloride, nor could the chlorine atom

be replaced by iodine by treatment with sodium iodide in acetone or in glacial acetic acid solution. No crystalline products were obtained when the 9(or 10)-chlororetene was nitrated or sulfonated. Attempts to obtain a monohydroxyretene from this monochlororetene by Grignard's reaction gave no identifiable products. Attempts to prepare a methylretene from this monochlororetene and methyl iodide with sodium in absolute ether were likewise unsuccessful. The foregoing indicates the stability of the C—Cl linkage and the unreactivity of the chlorine atom in 9(or 10)-chlororetene.

A mixture of mono-, di-, tri- and hexa-chlororetenes has recently been recommended for use in petroleum lubricants (148). Such a mixture of chlorinated retenes is added to a hydrocarbon lubricating base and has been proposed for high pressure or other purposes.

F. Retenequinone

Retenequinone (II) was first recognized as the quinone obtained by the oxidation of retene by Bamberger and Hooker (19, 20). Earlier workers had published confusing accounts of the nature of the quinone (15, 52, 54, 146).

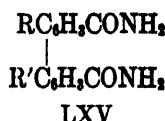
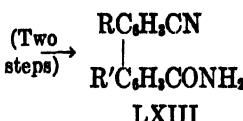
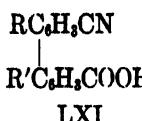
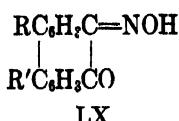
Retenequinone crystallizes in flat, orange prisms, m. p. 197°C. Valeur (142) has measured its heat of formation and its heat of combustion at constant volume and at constant pressure. The quinone is best prepared by oxidizing retene with chromic anhydride in acetic acid solution (22, page 116). Fieser and Young (61) have recommended that retenequinone be purified through its sodium bisulfite addition compound (53), which liberates the quinone upon acidification. Its sodium bisulfite addition compound is also described by Knesch (96).

Retenequinone reacts with *o*-phenylenediamine to give the quinoxaline (22, page 123). Its product with ethylenediamine has been described by Mason (107). With arylhydrazinesulfonic acids it yields dyes whose shades vary from orange-red to blue-red (8).

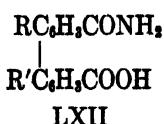
The monophenylhydrazone of retenequinone has been prepared by Bamberger and Grob (18). Its *p*-nitrophenylhydrazone has been reported by Li Man Cheung (103), and its semicarbazone and aminoguanidine derivatives by Heiduschka and Scheller (83). The dioxime of retenequinone has been described by Bamberger (16).

In an effort to cast light on the structure of retene, Lux (105) studied the Beckmann rearrangement of the mono $\ddot{\text{o}}$ xime of retenequinone (LX) (22, page 122). When treated with hydrogen chloride in a mixture of acetic acid and acetic anhydride, the oxime yielded chiefly a methyl-isopropylbiphenylenitrilocarboxylic acid (LXI), m. p. 112–114°C.,

together with a small amount of the corresponding methylisopropyl-diphenamic acid (LXII). The latter was also obtained by boiling the nitrilocarboxylic acid (LXI) with alcoholic potassium hydroxide. Thionyl chloride converted the nitrilocarboxylic acid (LXI) into the corresponding acid chloride, which furnished methylisopropylbiphenylenenitroamide (LXIII) when treated with ammonium hydroxide. Dilute alcoholic potassium hydroxide hydrolyzed the nitriloamide (LXIII) to the methylisopropylbiphenylenediamide (LXV). In one experiment Lux obtained the isomeric methylisopropylidiphenamic acid (LXIV) by the action of alcoholic potassium hydroxide on the nitriloamide (LXIII). The positions of the methyl and isopropyl groups relative to the other substituting groups in these compounds are uncertain. $R = (CH_3)_2CH-$ and $R' = CH_3-$, or *vice versa*, in formulas LX through LXV.



Retenequinone
monoöxime

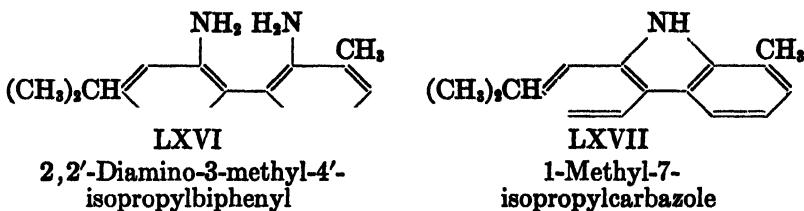


Werner and Piquet (151) obtained in poor yield a methylisopropylbiphenylenenitrilocarboxylic acid, m. p. 195°C., by the action of benzene-sulfonyl chloride on retenequinone monoöxime in pyridine solution. This nitrilocarboxylic acid was evidently isomeric with that (LXI) of Lux (105).

Bucher's brilliant work (39) establishing the structure of retene appeared almost simultaneously with Lux's (105) studies on the Beckmann rearrangement of retenequinone monoöxime. A few months later Lux published a more lengthy account of his work (106), together with the results of an extension of his experiments. By assuming that the iso-nitroso group ($=NOH$) in retenequinone monoöxime is at C₉ (1-methyl-7-isopropylphenanthraquinone-9-monoöxime), Lux assigned structures to the previously described compounds (LX through LXV). In formulas LX through LXV, Lux regarded R as $(CH_3)_2CH-$ and R' as CH_3- . His assumption was based on the observation that neither the nitrilocarboxylic acid (LXI) nor the diphenamic acid (LXII) could be esterified

by means of alcohol and concentrated sulfuric acid. From this he concluded that the carboxyl group in LXI and LXII must be ortho to both the methyl group (R') and the biphenyl linkage, and therefore could not be esterified because of steric hindrance. This additional evidence does not seem to be of sufficient weight to establish the structures of the compounds in question (LX through LXV), and the definition of R as $(CH_3)_2CH-$ and R' as CH_3- must be regarded as provisional.

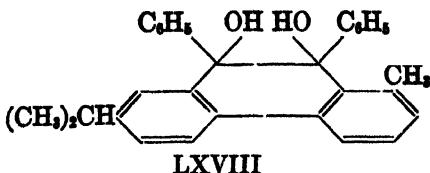
The Hofmann degradation of the diamide (LXV) yielded the expected 2,2'-diamino-3-methyl-4'-isopropylbiphenyl (LXVI). When Lux tetratized this diamino compound and decomposed the tetrazonium salt thus produced under conditions usually employed for hydrocarbon formation, he did not obtain the expected 3-methyl-4'-isopropylbiphenyl. Instead, 1-methyl-7-isopropylcarbazole (LXVII) was secured.



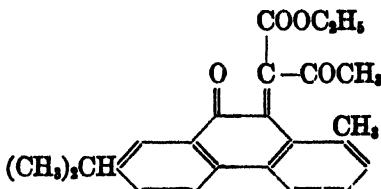
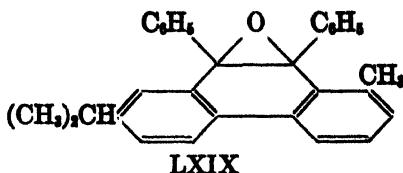
Bamberger and Hooker (22, page 145) have reported the behavior of retenequinone on further oxidation. Potassium permanganate oxidation yielded 7-(α -hydroxyisopropyl)fluorenone-1-carboxylic acid (VI), fluorenone-1,7-dicarboxylic acid (VII), 1-methyl-7-isopropylfluorenone, oxalic acid, and other products. The quinone was reduced to retene by phosphorus and hydriodic acid.

Retenequinone yields reteneglycolic acid (III) when boiled with strong aqueous alkali (22, page 132) or with 10 per cent methyl alcoholic potassium hydroxide (104). By fusion of the quinone with potassium hydroxide and lead peroxide, Lux (104) obtained 4'-isopropyl-3-methylbiphenyl-2-carboxylic acid (IV).

By the action of suitable organomagnesium halogen compounds on retenequinone, Heiduschka and Grimm (80) have prepared a number of 9,10-dihydroxy-9,10-diaryl(or dialkyl)-9,10-dihydroretene derivatives. 9,10-Dihydroxy-9,10-diphenyl-9,10-dihydroretene (LXVIII) yielded an anhydride (LXIX) when treated with acetyl chloride. When the dihydroxy compound (LXVIII) was heated with zinc dust, 9,10-diphenylretene was formed. Reduction of LXVIII with hydriodic acid and red phosphorus yielded a 9,10-diphenylhexahydroretene, and finally a retene-dodecahydride (101).



9,10-Dihydroxy-9,10-diphenyl-
9,10-dihydroretene



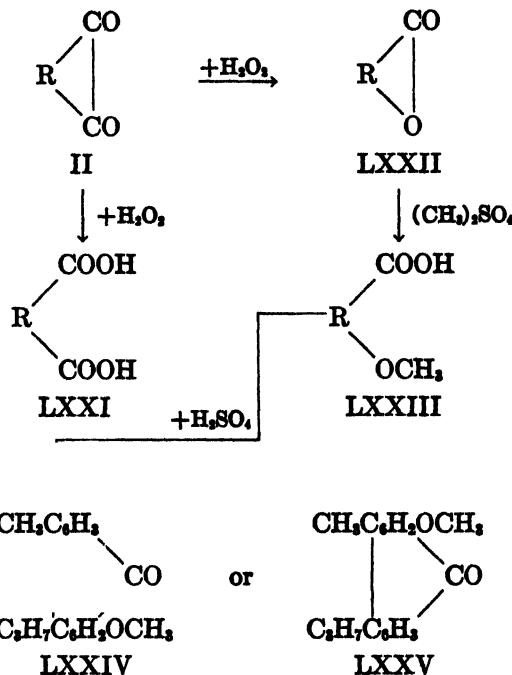
LXX
Ethyl retroxyleneacetoacetate

Retenequinone condenses with ketones in the presence of aqueous or alcoholic potassium hydroxide (81). Heiduschka and Khudadad have prepared a wide variety of such condensation products. The structures of the latter have not been absolutely established. When retenequinone is boiled with acetoacetic ester in alcoholic solution in the presence of a little piperidine or potassium hydroxide, ethyl retroxyleneacetoacetate (LXX) is formed (82). A number of reactions and derivatives of this product have been investigated.

Adelson, Hasselstrom, and Bogert (6) have studied the oxidation of retenequinone in glacial acetic acid by means of 30 per cent hydrogen peroxide. Two substances were obtained, retenediphenic acid (LXXI) and a small amount of a lactone by-product. The latter, a methylisopropylhydroxybiphenylcarboxylic lactone (LXXII), was identified as such by conversion into the methylisopropylmethoxybiphenylcarboxylic acid (LXXIII), and this in turn into the methylisopropylmethoxyfluorenone-9 (LXXIV or LXXV). In the formulas shown at top of next

page, the residue $\text{CH}_3\text{C}_6\text{H}_3\text{C}_6\text{H}_4\text{C}_2\text{H}_5$, is represented by R.

Reolediphenic acid was found to behave as a true diphenic acid and to resemble phthalic acid in the formation of artificial resins with glycerol or borneol. The acid had previously been obtained by Fogelberg (63) as a syrupy oxidation product when retenequinone was oxidized with hydrogen peroxide. The desired acid could be obtained in crystalline condition only by conversion to the anhydride, followed by purification and hydrolysis of the latter.



G. 6-Acetylretene

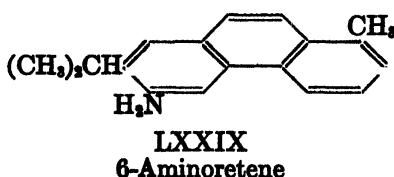
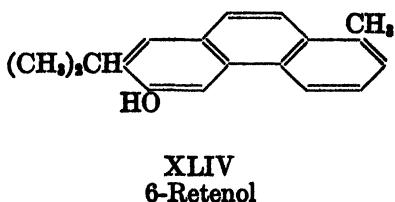
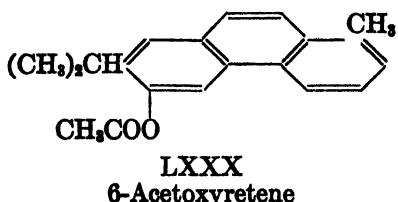
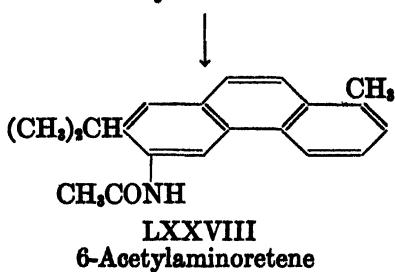
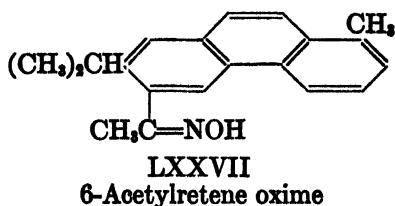
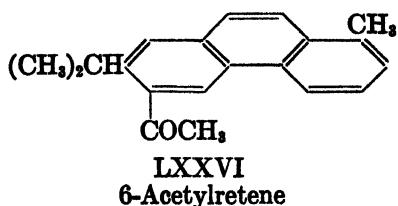
Of all the derivatives of retene which have been prepared to date one of the most important is 6-acetylretene (LXXVI). As a synthetic intermediate, 6-acetylretene has proven to be of inestimable value. Its worth as a point of departure in proof of structure work is equally great, and it can be stated without reservation that this important compound has been very instrumental in guiding the course of much synthetic activity in progress during the past year.

6-Acetylretene (1, 32) is the only acetylated retene of established structure that is known at present. The acetylation of 9,10-dihydroretene (114) has been investigated by Nyman. Freund and Fleischer (65) have reported a "retenediethylindandione" from retene, diethylmalonyl chloride, and aluminum chloride. Neither Freund and Fleischer nor Nyman have proven the structure of their products. Another acetylated compound of unestablished structure was obtained by Perrier (117). He reported that the compound $(\text{C}_6\text{H}_5\text{COCl})_2\text{Al}_2\text{Cl}_6$, obtained by heating benzoyl chloride with aluminum chloride in carbon bisulfide, reacted with retene to form a compound, $(\text{C}_6\text{H}_5\text{COC}_1\text{H}_{17})_2\text{Al}_2\text{Cl}_6$.

The preparation of a condensation product of retene with formaldehyde, said to be suitable for tanning, has been patented (14).

The synthesis of 6-acetylretene was developed by Bogert and Hasselstrom (32). It involved the Friedel-Crafts reaction between retene, acetyl chloride, and aluminum chloride in carbon bisulfide solution. The yield and technique of the synthesis were improved by Adelson and Bogert (1), who also established the structure of the compound by a sequence of reactions discussed in the following paragraph. Recently a patent (147) has appeared in which nitrobenzene is used as a solvent in the acetylation of retene.

6-Acetylretene oxime (LXXVII) (32) was rearranged to 6-acetylaminoretene (LXXVIII) (1) by means of the well-known Beckmann rearrangement. The latter compound was hydrolyzed by alcoholic potassium hydroxide, and the 6-aminoretene (LXXIX) thus formed was converted into 6-retenol (XLIV) through the agency of the diazo reaction. 6-



Retenol (XLIV) was acetylated, and both the acetylated product, 6-acetoxyretene (LXXX), and the parent phenol were found to be identical with authentic specimens of the materials prepared by other means (61, 98).

The Beckmann rearrangement of a ketoxime has been applied very recently and independently by several investigators to the preparation of phenanthryl amines (1, 12, 112).

As in acetophenone, the terminal group of 6-acetylretene was very reactive and condensed smoothly with various aldehydes to chalcone types, some of which possessed tinctorial properties (32). Bromination of 6-acetylretene gave 6- ω -bromoacetylretene and 6- ω -dibromoacetylretene (5). 6-Acetylretene was also found to undergo the Reformatsky reaction (3, 4) (see below).

By the method of Wolff (153) 6-acetylretene semicarbazone was easily reduced to 6-ethylretene (32).

The oxidation of 6-acetylretene to retene-6-carboxylic acid (1, 33) has been discussed above.

H. Polynuclear compounds

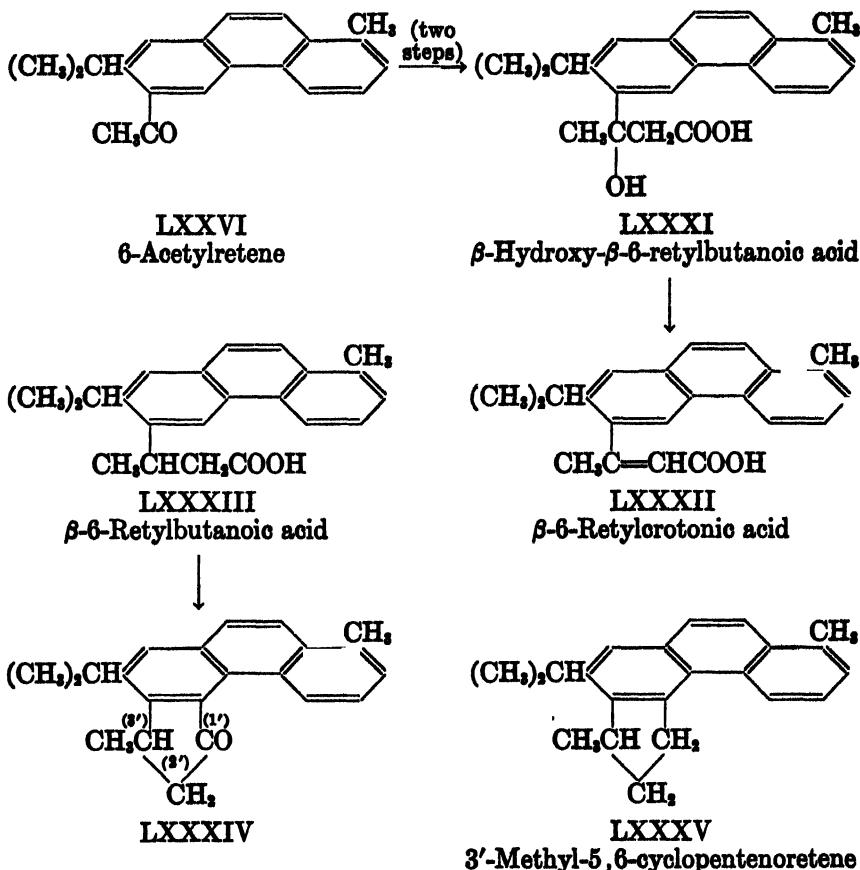
By fusion with selenium the so-called "Diels' hydrocarbon" (3'-methyl-1,2-cyclopentenophenanthrene) has been obtained from such biologically interesting and chemically important compounds as cholesterol, ergosterol, lumisterol, isopyrovitamin, pyrocalciferol, cholic acid, strophantidin, gitogenin, anhydrouzarigenin, sarsasapogenin, and pseudobufotalin.

Since retene is itself a natural product containing the methyl and isopropyl groups, two radicals which are frequently found associated in natural products and which seem to have some significant function in nature, Adelson and Bogert (3, 4, 5) have launched a program involving the synthesis of polynuclear hydrocarbons derived from retene. This has thus far included the synthesis of 3'-methyl-5,6-cyclopentenoretene (LXXXV) (3, 4) and 5,6-benzoretene (XCII) (5). The former is of significance because it is a methyl isopropyl homolog of the Diels' hydrocarbon; the latter, 5,6-benzoretene (XCII), is of considerable interest because it is a methyl isopropyl homolog of 3,4-benzophenanthrene (42), which possesses considerable carcinogenic activity and which is the simplest carcinogenic substance yet encountered (27).

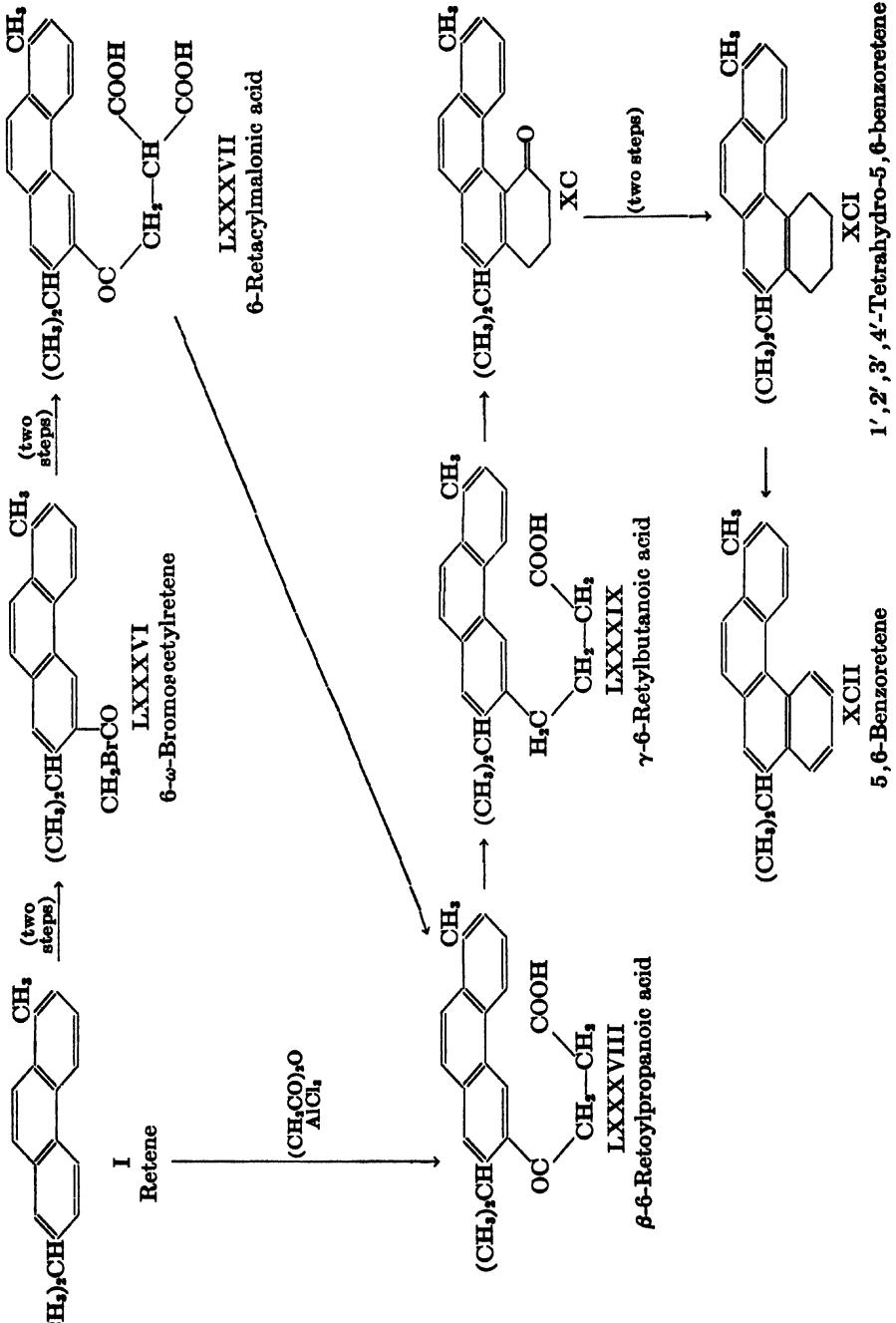
(1) 3'-Methyl-5,6-cyclopentenoretene

This hydrocarbon has been synthesized by Adelson and Bogert (3, 4). The Reformatsky reaction involving 6-acetylretene (LXXVI), ethyl bromoacetate, and zinc dust yielded a hydroxy ester which was hydrolyzed to β -hydroxy- β -6-retylbutanoic acid (LXXXI). The hydroxy acid

(LXXXI) was dehydrated by means of acetic anhydride to β -6-retylcrotonic acid (LXXXII), which was then reduced to β -6-retylbutanoic acid (LXXXIII). Cyclization of the acid chloride of the latter through the agency of aluminum chloride in benzene solution gave 1'-keto-3'-methyl-5,6-cyclopentenoretene (LXXXIV). This cyclic ketone was converted into 3'-methyl-5,6-cyclopentenoretene (LXXXV) by means of the Clemmensen reduction.



Inasmuch as 6-acetylretene was used as the starting material, the structure of 3'-methyl-5,6-cyclopentenoretene (LXXXV) was defined at once. Since C₇ was occupied by the isopropyl group, only C₆ was available for ring closure. The same reasoning applies to 5,6-benzoretene, which is discussed below.

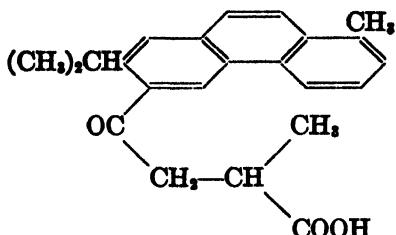


(2) 5,6-Benzoretene

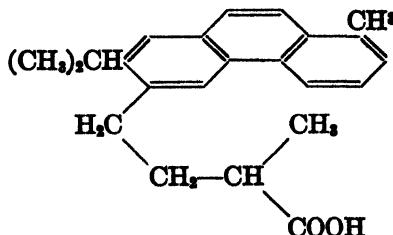
5,6-Benzoretene (XCII) has been synthesized by Adelson and Bogert (5) by a series of steps, starting with retene (I). The bromination of 6-acetylretene (LXXVI) gave 6- ω -bromoacetylretene (LXXXVI) and a small amount of 6- ω -dibromoacetylretene. The monobromo compound (LXXXVI) was condensed with sodiomalonic ester, and the resulting retacylmalonic ester was hydrolyzed to the corresponding 6-retacylmalonic acid (LXXXVII). Decarboxylation of the latter yielded β -6-retarylpropanoic acid (LXXXVIII), which was identical with the keto acid obtained by the interaction of succinic anhydride, retene, and aluminum chloride in benzene solution. The keto acid (LXXXVIII) was reduced by the modified Clemmensen method to γ -6-retylbutanoic acid (LXXXIX). Treatment of the latter with phosphorus pentachloride followed by aluminum chloride gave 1'-keto-1', 2', 3', 4'-tetrahydro-5,6-benzoretene (XC); this keto compound was also obtained by the action of anhydrous stannic chloride on the acid (LXXXIX). The semicarbazone of this cyclic ketone (XC) was reduced by the method of Wolff (153) to 1', 2', 3', 4'-tetrahydro-5,6-benzoretene (XCI). The latter was smoothly dehydrogenated by sulfur to 5,6-benzoretene (XCII).

When retene was treated with pyrotartaric anhydride in the presence of aluminum chloride, α -methyl- β -6-retarylpropanoic acid (XCIII) resulted (5). Its structure was established by its synthesis from 6- ω -bromoacetylretene (LXXXVI) and sodio-methylmalonic ester in a manner similar to that described above for β -6-retarylpropanoic acid (LXXXVIII). The α -methyl- β -6-retarylpropanoic acid (XCIII) from both syntheses was identical. It was reduced by means of the modified Clemmensen method to α -methyl- γ -6-retylbutanoic acid (XCIV), which was cyclized by sulfuric acid to 1'-keto-2'-methyl-1', 2', 3', 4'-tetrahydro-5,6-benzoretene (XCV). The ultimate goal is 2'-methyl-5,6-benzoretene.

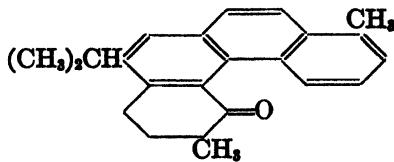
In the course of experimental work on the synthesis of 5,6-benzoretene (XCII), Adelson and Bogert (5) investigated an alternate route leading to this hydrocarbon. The method involved was a modification of the von Auwers and Möller (11) adaptation of the Bougault synthesis (36) as employed by Fieser and Hershberg (58). Methyl γ -6-retylbutanoate (XCVI) was condensed with ethyl oxalate in the presence of sodium ethylate. The sodium salt of the reaction product was treated with 10 per cent sulfuric acid to yield the free keto ester (XCVII), which was subsequently cyclized by means of 3', 4'-dihydro-5,6-benzoretene-1', 2'-dicarboxylic acid anhydride (XCVIII). Dehydrogenation of the latter with sulfur gave 5,6-benzoretene-1', 2'-dicarboxylic acid anhydride (XCIX), which could not be successfully decarboxylated.



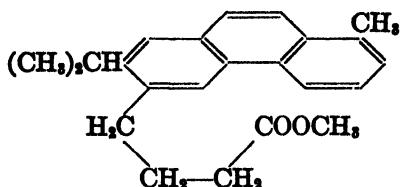
XCIII
 α -Methyl- β -6-retarylpropanoic acid



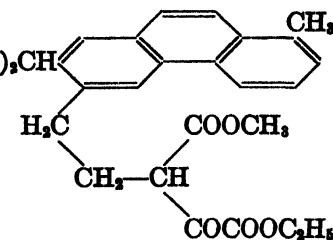
XCIV
 α -Methyl- γ -6-retylbutanoic acid



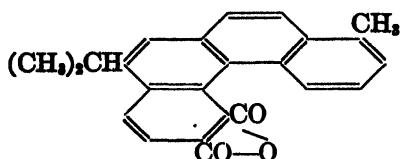
XCV
1'-Keto-2'-methyl-1',2',3',4'-tetrahydro-5,6-benzoretene



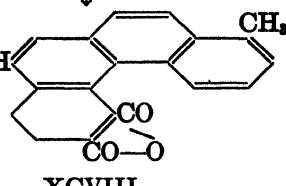
XCVI
Methyl γ -6-retylbutanoate



XCVII



XCIX
5,6-Benzoretene-1',2'-dicarboxylic
acid anhydride



XCVIII

VIII. CONCLUSION

A. Relation of retene to the terpenes

The sulfur dehydrogenation method first employed by Vesterberg (143) in obtaining retene from abietic acid has been of the greatest importance

in the investigation of other natural products. This method was highly instrumental in the determination of the structures of the dicyclic sesquiterpenes ($C_{15}H_{34}$), which have been found to yield either cadalene ($C_{15}H_{18}$; 1,6-dimethyl-4-isopropylnaphthalene) or eudalene ($C_{14}H_{16}$; 1-methyl-7-isopropylnaphthalene) (119, 123, 125). These two hydrocarbons, cadalene and eudalene, thus occupy key positions in relating the dicyclic sesquiterpenes and the naphthalene series.

It has also been shown that the monocyclic terpenes, such as terpinene and limonene, give *p*-cymene upon dehydrogenation with sulfur (125). *p*-Cymene stands in the same relationship to the monocyclic terpenes and the benzene series as do eudalene and cadalene in the case of the dicyclic sesquiterpenes and the naphthalene series.

It is highly probable that retene occupies a similar position in phenanthrene chemistry in relation to the diterpenes. Abietic acid, which may be regarded as a diterpene type, yields retene when heated with sulfur. Unfortunately, in this instance the sulfur dehydrogenation has not proven to be as fruitful in elucidating the structure of the parent compound as in the case of the sesquiterpenes. It is possible that the recent work on hinokiol (92, 154) may throw additional light on the relationship between retene and diterpenes.

B. Comparison of retene and phenanthrene

Retene does not generally behave as one would expect a methyl isopropyl homolog of phenanthrene to behave.

While retene itself yields only resinous products upon nitration, the nitration of phenanthrene in glacial acetic acid solution with diacetyl ortho nitric acid yields 9-nitrophenanthrene as the chief product, together with smaller amounts of the 2- and 4-nitro derivatives and a very small amount of the 3-nitro isomer (95, 130).

Only one amino derivative of retene is known, 6-aminoretene (1). Aminophenanthrenes, prepared as in the case of 6-aminoretene by hydrolysis of the product of the Beckmann rearrangement of the appropriate acetylphenanthrene oxime, have been synthesized where the amino group is at C₁, C₂, C₈, and C₉ (12, 112).

No systematic work has been done on the halogen derivatives of retene. Perhaps the only definitely known retyl halide is 9(or 10)-chlororetene (98). 9-Bromophenanthrene is an easily accessible compound. Other phenanthryl halides (1-, 2-, and 3-chloro, -bromo, and -iodo) have been prepared by diazotization of the appropriate amines (13, 60).

Sulfonation of retene gives acids with the sulfonic acid group at C₂ and C₈ (61, 98). The isomers formed in largest amount on sulfonating phenanthrene are the 2-sulfonic and the 3-sulfonic acids (the latter may also

be called the 6-sulfonic acid); the 1-sulfonic and the 9-sulfonic acid are also known (56, 86, 129, 150).

The acetylation of retene by means of the Friedel-Crafts reaction, using carbon bisulfide or nitrobenzene as the solvent, gives 6-acetylretene (1, 32, 147). Under similar conditions phenanthrene is converted mainly into resinous products. If nitrobenzene is used as the solvent, 2- and 3-acetylphenanthrenes are obtained in good yield (110).

In the presence of aluminum chloride, succinic anhydride and pyrotartaric anhydride condense smoothly with retene in benzene solution and give exclusively products containing substituents in the 6-position (5). When condensed with phenanthrene in nitrobenzene solution, these anhydrides yield principally the C₃ compounds (43, 79). In general, the results with phenanthrene indicate that C₃ (which may also be regarded as C₆) is particularly favored in the Friedel-Crafts reaction.

6-Retenecarboxylic acid is most conveniently prepared by the action of alkaline iodine-potassium iodide solution on 6-acetylretene (1). The 2- and 3-carboxylic acids of phenanthrene are formed in quantitative yield by the action of aqueous sodium hypochlorite solution on the appropriate acetyl compound (110, 111).

No retylaldehydes or retynitriles are known. Distillation of potassium retene-6-sulfonate with potassium ferrocyanide in an attempt to prepare a nitrile gave no good product (61).

The readily available phenanthrols, 2- and 3-phenanthrols, are prepared from the sulfonates (56). 2- and 6-Retenols are synthesized in similar fashion (61, 98). While other phenanthrols are known, namely, 1-phenanthrol (56, 109, 135), 4-phenanthrol (28), and 9-phenanthrol (40, 44, 59, 71, 131, 134), the corresponding possible retenols are not known.

From the foregoing it is evident that substitution in the retene molecule takes place with far less formation of isomers than in the case of phenanthrene. Fieser and Young (61) state: "The alkyl groups (of retene) serve only to promote substitution at an adjacent position and thus to hinder the formation of a large number of isomers."

IX. REFERENCES

- (1) ADELSON AND BOGERT: J. Am. Chem. Soc. **58**, 653 (1936).
- (2) ADELSON AND BOGERT: J. Am. Chem. Soc. **58**, 2236 (1936).
- (3) ADELSON AND BOGERT: J. Am. Chem. Soc. **59**, 399 (1937).
- (4) ADELSON AND BOGERT: Proc. Natl. Acad. Sci. U. S. **23**, 117 (1937).
- (5) ADELSON AND BOGERT: J. Am. Chem. Soc. **59**, 1776 (1937).
- (6) ADELSON, HASSELSTROM, AND BOGERT: J. Am. Chem. Soc. **58**, 871 (1936).
- (7) Aktien Gesellschaft für chemische Industrie in Rheinau: English patent 13,366 (1887); German patent 43,802 (1887); Ber. **21R**, 553 (1888).

- (8) Aktien Gesellschaft für chemische Industrie in Rheinau: English patent 4217 (1888); German patent 46,746; Frdl. 2, 748.
- (9) ARNOT: German patent 315,623 (1919); Chem. Zentr. 1920, II, 188.
- (10) ARNOT: British patent 149,354 (1920); Chem. Zentr. 1921, II, 37.
- (11) VON AUWERS AND MÖLLER: J. prakt. Chem. 109, 124 (1925).
- (12) BACHMANN AND BOATNER: J. Am. Chem. Soc. 58, 857, 2097 (1936).
- (13) BACHMANN AND BOATNER: J. Am. Chem. Soc. 58, 857, 2194 (1936).
- (14) Badische Anilin und Soda Fabrik: German patent 290,965 (1917).
- (15) BAMBERGER: Ber. 17, 453 (1884).
- (16) BAMBERGER: Ber. 18, 81 (1885).
- (17) BAMBERGER: Ber. 22, 635 (1889).
- (18) BAMBERGER AND GROB: Ber. 34, 533 (1901).
- (19) BAMBERGER AND HOOKER: Ber. 18, 1024 (1885).
- (20) BAMBERGER AND HOOKER: Ber. 18, 1030 (1885).
- (21) BAMBERGER AND HOOKER: Ber. 18, 1750 (1885).
- (22) BAMBERGER AND HOOKER: Ann. 229, 102 (1885).
- (23) BAMBERGER AND LODTER: Ber. 20, 3076 (1887).
- (24) BAMBERGER AND STRASSER: Ber. 22, 3361 (1889).
- (25) BARDHAN AND SENGUPTA: J. Chem. Soc. 1932, 2521.
- (26) BARDHAN AND SENGUPTA: J. Chem. Soc. 1932, 2798.
- (27) BARRY, COOK, HASLEWOOD, HEWETT, HIEGER, AND KENNAWAY: Proc. Roy. Soc. (London) B117, 351 (1935).
- (28) BEHREND AND LUDEWIG: Ann. 379, 351 (1911).
- (29) BERTHELOT: Bull. soc. chim. [2] 8, 389 (1867).
- (30) BERTHELOT AND RECOURA: Ann. chim. phys. [6] 13, 298 (1888).
- (31) BERTHELOT AND VIEILLE: Ann. chim. phys. [6] 10, 447 (1887).
- (32) BOGERT AND HASSELSTROM: J. Am. Chem. Soc. 53, 3462 (1931).
- (33) BOGERT AND HASSELSTROM: Proc. Natl. Acad. Sci. U. S. 18, 417 (1932).
- (34) BOGERT AND HASSELSTROM: J. Am. Chem. Soc. 56, 983 (1934).
- (35) BOGERT AND STERLING: Science 87, 196, 234 (1938).
- (36) BOUGAULT: Compt. rend. 159, 745 (1915).
- (37) BOYEN: Chem.-Ztg. 11, 1297 (1887).
- (38) BROMIES: Ann. 37, 304 (1841).
- (39) BUCHER: J. Am. Chem. Soc. 32, 374 (1910).
- (40) CHATTERJEE: J. Indian Chem. Soc. 12, 591 (1935).
- (41) CHILESOTTI: Gazz. chim. ital. 30, I, 159 (1900).
- (42) COOK: J. Chem. Soc. 1931, 2524.
- (43) COOK AND HASLEWOOD: J. Chem. Soc. 1934, 428.
- (44) COOK, HEWETT, AND LAWRENCE: J. Chem. Soc. 1936, 71.
- (45) CZERNY: Bul. soc. chim. România 7, 91 (1925); Chem. Abstracts 20, 1320 (1926).
- (46) DIELS AND KARSTENS: Ber. 60, 2323 (1927).
- (47) DIMROTH: Ann. 446, 97 (1926).
- (48) DIMROTH AND FAUST: Ber. 54, 3020 (1921).
- (49) DITZ: Chem.-Ztg. 31, 486 (1907).
- (50) EASTERFIELD AND BAGLEY: J. Chem. Soc. 85, 1238 (1904).
- (51) EKSTRAND: Bull. soc. chim. [2] 24, 55 (1875).
- (52) EKSTRAND: Ber. 9, 855 (1876).
- (53) EKSTRAND: Ann. 185, 75 (1877).
- (54) EKSTRAND: Ber. 17, 692 (1884).

- (55) FEHLING: Ann. **106**, 388 (1858).
(56) FIESER: J. Am. Chem. Soc. **51**, 2460 (1929).
(57) FIESER: J. Am. Chem. Soc. **51**, 3101 (1929).
(58) FIESER AND HERSHBERG: J. Am. Chem. Soc. **57**, 1851 (1935).
(59) FIESER, JACOBSEN, AND PRICE: J. Am. Chem. Soc. **58**, 2163 (1936).
(60) FIESER AND PRICE: J. Am. Chem. Soc. **58**, 1838 (1936).
(61) FIESER AND YOUNG: J. Am. Chem. Soc. **53**, 4120 (1931).
(62) FITTIG AND OSTERMAYER: Ber. **5**, 933 (1872); Ann. **166**, 361 (1873).
(63) FOGLERBERG: Ann. Acad. Sci. Fennicae **A29**, No. 4, 3-7 (Komppa Festschrift) (1927); Chem. Zentr. **98**, II, 2299 (1927).
(64) FORTNER: Monatsh. **25**, 443 (1904).
(65) FREUND AND FLEISCHER: Ann. **373**, 291 (1910).
(66) FRITZSCHE: Jahresber. Chem. **1858**, 440.
(67) FRITZSCHE: J. prakt. Chem. **75**, 281 (1858).
(68) FRITZSCHE: Ann. **109**, 250 (1859).
(69) FRITZSCHE: J. prakt. Chem. **82**, 321 (1861).
(70) FRITZSCHE: J. prakt. Chem. **82**, 335 (1861).
(71) GOLDSCHMIDT, VOGT, AND BREDIG: Ann. **445**, 135 (1925).
(72) GOLDSTEIN: Chem. Zentr. **1911**, II, 342, 586.
(73) HANSEN: Ber. **42**, 214 (1909).
(74) HASSELSTROM AND BOGERT: J. Am. Chem. Soc. **57**, 1579 (1935).
(75) HASSELSTROM AND HULL: United States patent 2,095,548 (1937).
(76) HAWORTH: J. Chem. Soc. **1932**, 1125.
(77) HAWORTH: J. Chem. Soc. **1932**, 2717.
(78) HAWORTH, LETSKY, AND MAVIN: J. Chem. Soc. **1932**, 1784.
(79) HAWORTH AND MAVIN: J. Chem. Soc. **1933**, 1012.
(80) HEIDUSCHKA AND GRIMM: Arch. Pharm. **250**, 33 (1912); Chem. Zentr. **1912**, I, 727.
(81) HEIDUSCHKA AND KHUDADAD: Arch. Pharm. **261**, 401 (1913); Chem. Zentr. **1913**, II, 2043.
(82) HEIDUSCHKA AND KHUDADAD: Arch. Pharm. **261**, 682 (1913); Chem. Abstracts **8**, 2699 (1914).
(83) HEIDUSCHKA AND SCHELLEK: Arch. Pharm. **248**, 89 (1910); Chem. Zentr. **1910**, I, 1974.
(84) HELL: Ber. **22**, 498 (1889).
(85) HENKE AND CASTAO: United States patent 1,881,565 (1932).
(86) IOFFÉ: J. Gen. Chem. (U. S. S. R.) **3**, 448 (1933).
(87) IPATIEV: Ber. **42**, 2092 (1909).
(88) JEFREMOW: J. Russ. Phys. Chem. Soc. **50**, 403 (1918); Chem. Zentr. **1923**, III, 380.
(89) JEFREMOW: J. Russ. Phys. Chem. Soc. **50**, 436 (1918); Chem. Zentr. **1923**, III, 380.
(90) JEFREMOW: J. Russ. Phys. Chem. Soc. **50**, 451 (1918); Chem. Zentr. **1923**, III, 381.
(91) JEFREMOW: J. Russ. Phys. Chem. Soc. **51**, 379 (1919); Chem. Zentr. **1923**, III, 770.
(92) KEIMATSU AND ISHIGURO: J. Pharm. Soc. Japan **55**, 186 (1935) (in German, p. 45); Chem. Abstracts **29**, 7323 (1935); Chem. Zentr. **1935**, II, 3664.
(93) KEIMATSU, ISHIGURO, AND SUMI: J. Pharm. Soc. Japan **56**, 119 (1936); Chem. Zentr. **1937**, I, 2188.

- (94) KELBE: Chem. Zentr. 1887, 1504.
 (95) KEYES AND BROOKER: J. Am. Chem. Soc. 59, 74 (1937).
 (96) KNESCH: German patent 151,981 (1903).
 (97) KOMPPA AND FOGELBERG: J. Am. Chem. Soc. 54, 2900 (1932).
 (98) KOMPPA AND WAHLFORSS: J. Am. Chem. Soc. 52, 5009 (1930).
 (99) KRAFT AND WEILANDT: Ber. 29, 2241 (1896).
 (100) KRÄMER: Ber. 38, 647 (1903).
 (101) LIEBERMANN AND SPIEGEL: Ber. 22, 779 (1889).
 (102) LIEBERMANN AND ZSUPPA: Ber. 44, 857 (1911).
 (103) LI MAN CHEUNG: Bull. inst. pin. 1929, 108, 189, 188, 215; Chem. Abstracts 23, 3467, 4464 (1929); Chem. Zentr. 1929, II, 1528.
 (104) LUX: Monatsh. 29, 763 (1908).
 (105) LUX: Ber. 43, 688 (1910).
 (106) LUX: Monatsh. 31, 939 (1910).
 (107) MASON: J. Chem. Soc. 63, 1288 (1893).
 (108) MORGENTHORN: Monatsh. 31, 292 (1910).
 (109) MOSSETTIG AND BURGER: J. Am. Chem. Soc. 57, 2189 (1935).
 (110) MOSSETTIG AND VAN DE KAMP: J. Am. Chem. Soc. 52, 3704 (1930).
 (111) MOSSETTIG AND VAN DE KAMP: J. Am. Chem. Soc. 55, 2995 (1933).
 (112) MOSSETTIG AND KRUEGER: J. Am. Chem. Soc. 58, 1811 (1936).
 (113) MUNCIE: United States patent 2,069,896 (1937).
 (114) NYMAN: Ann. Acad. Sci. Fenniae A41, No. 5, 80 pp. (1934); Chem. Abstracts 30, 2958 (1936).
 (115) ORLOV: J. Russ. Phys. Chem. Soc. (Chemical part) 60, No. 9, 1447 (1928); Chem. Abstracts 23, 2174 (1929).
 (116) PERLMAN, DAVIDSON, AND BOGERT: J. Org. Chem. 1, 288, 300 (1936).
 (117) PERRIER: Compt. rend. 116, 1298 (1893).
 (118) POCHETTINO: Atti. accad. Lincei [5] 18, II, 360 (1909); Chem. Zentr. 1910, I, 721.
 (119) RUZICKA: Fortschr. Chem. Physik physik. Chem. 19, Heft 5, 1 (1928).
 (120) RUZICKA, BALAS, AND SCHINZ: Helv. Chim. Acta 6, 692 (1923).
 (121) RUZICKA, DE GRAAF, AND HOSKING: Helv. Chim. Acta 14, 233 (1931).
 (122) RUZICKA, DE GRAAF, AND MÜLLER: Helv. Chim. Acta 15, 1800 (1932).
 (123) RUZICKA AND MEYER: Helv. Chim. Acta 4, 505 (1921).
 (124) RUZICKA AND MEYER: Helv. Chim. Acta 5, 581 (1922).
 (125) RUZICKA, MEYER, AND MINGAZZINI: Helv. Chim. Acta 5, 345 (1922).
 (126) RUZICKA AND WALDMANN: Helv. Chim. Acta 15, 907 (1932).
 (127) RUZICKA AND WALDMANN: Helv. Chim. Acta 16, 842 (1933).
 (128) RUZICKA AND WALDMANN: Helv. Chim. Acta 18, 611 (1935).
 (129) SANDQVIST: Ann. 392, 76 (1912).
 (130) SCHMIDT AND HEINLE: Ber. 44, 1488 (1911).
 (131) SCHMIDT AND LUMPP: Ber. 41, 4215 (1908).
 (132) SCHULTZ: Ber. 11, 215 (1878); 12, 235 (1879).
 (133) SCHULTZ: Ann. 359, 129 (1908).
 (134) SHERWOOD, SHORT, AND WOODCOCK: J. Chem. Soc. 1936, 322.
 (135) SHOEMAKER AND GUTHRIE: J. Chem. Soc. 1928, 2332.
 (136) SPIEGEL: Ber. 22, 3369 (1889).
 (137) STORMANN: Z. physik. Chem. 6, 339 (1890).
 (138) STRAUS: Ber. 51, 1466 (1918).
 (139) SUDBOROUGH: J. Chem. Soc. 109, 1344 (1916).

- (140) TROMMSDORF: Ann. **21**, 126 (1837).
- (141) TSCHIRCH AND STUDER: Dissertation, Bern, 1903.
- (142) VALEUR: Compt. rend. **126**, 1148 (1898).
- (143) VESTERBERG: Ber. **36**, 4200 (1903).
- (144) VIRTANEN: Ber. **53**, 1885 (1920).
- (145) VIRTANEN: Ann. **424**, 150 (1921).
- (146) WAHLFORSS: Jahresber. Chem. **1869**, 501.
- (147) WAHLFORSS AND GOLDBLATT (to the Glidden Company): United States patent 2,054,107 (September 15, 1936).
- (148) WAHLFORSS, JOHNSON, AND KAY: British patent 431,508 (1935).
- (149) WERNER AND FREY: Ann. **321**, 298 (1902).
- (150) WERNER, FREY, KUNZ, J., KUNZ, M., LÖWENSTEIN, REKNER, AND WACK: Ann. **321**, 248 (1902).
- (151) WERNER AND PIQUET: Ber. **37**, 4295 (1904).
- (152) WIEDEMANN AND SCHMIDT: Ann. phys. chim. [2] **56**, 18 (1895); J. Chem. Soc. **70A**, 86 (1896).
- (153) WOLFF: Ann. **394**, 86 (1912).
- (154) YOSHIKI AND ISHIGURO: J. Pharm. Soc. Japan **53**, 73 (1933) (in German, p. 12); Chem. Abstracts **27**, 4530 (1933); Chem. Zentr. **1933**, I, 3201.

A SYMPOSIUM ON THE PHYSICAL CHEMISTRY OF THE PROTEINS¹

INTRODUCTION TO THE SYMPOSIUM

J. W. WILLIAMS

Department of Chemistry, University of Wisconsin, Madison, Wisconsin

Received January 30, 1939

At the semi-annual meetings of the American Chemical Society and on other occasions some of the Divisions of the Society, either alone or in collaboration with others, have organized symposia, the purpose of which has been the consideration of the advances in research in the several fields of their scope and the presentation of authoritative and critical summaries of those subjects which are important for the development of pure and applied chemistry. It may be said of the proteins, more than of any other group of compounds, that they are the essence of life, performing as they do in living organisms many important and diverse functions. In recent years data having to do with the nature and behavior of the proteins obtained by physical-chemical observations as distinct from organic-chemical and biochemical methods have been accumulating rapidly, so that it seems a suitable time to attempt to coördinate the knowledge that has become available. The Divisions of Colloid Chemistry and of Physical and Inorganic Chemistry have united their efforts with this end in view.

To study proteins as macromolecules, it is required first of all to learn about their mass, shape, and structure. This is accomplished by the use of experimental procedures such as surface films, surface potential differences, sedimentation, diffusion, osmotic pressure, electrophoresis, viscosity, double refraction of flow, dispersion of dielectric constant, etc. The results arrived at through work along these lines are summarized by the participants in this Symposium.

The chemist is interested in these substances because they have many, if not all, of the properties of molecules. The organic chemist considers their composition, which is derivable from a study of the hydrolytic decompositions. On the other hand, the physical and colloid chemist is

¹ This Symposium was held by the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the Ninety-sixth Meeting of the American Chemical Society, held in Milwaukee, Wisconsin, September, 1938.

concerned with the molecular weight, molecular form, and solubility relationships, or in a few words, the behavior of these substances as macromolecules. In the reports to follow, the chief concern is the structure and nature of the proteins regarded as physical-chemical units and not as organic compounds.

In the opening paper Langmuir and Schaefer give an account of their recent important work with protein films. After a brief description of the methods used in the formation of the films, their structure is discussed in terms of such measureable properties as thickness, compressibility, insolubility, viscosity, plasticity, and anisotropy after shear or linear compression. Possible explanations are given for certain of these properties in terms of the Wrinch cyclol theory. In the concluding sections there are described some new aspects of the specific reactivities of urease, catalase, and pepsin monolayers.

The extraordinary importance of solubility is universally recognized in chemistry. Recent progress has resulted because investigators are learning how to take into account forces of interaction between ions, dipole ions, and solvent molecules. Experimental and theoretical contributions of distinct merit in these directions have been made by the authors of the next two reports. On the experimental side of the subject, Cohn and his associates have provided data which are indispensable for the development of our subject. In the present report we find a discussion of the parameters in terms of which the structure of protein molecules may be described. Thus, detailed consideration is given to methods by which information is gained regarding the size and shape of protein molecules and by which the number and distribution of the electrically charged groups are determined. Extremely useful tables of certain physical constants for a number of proteins are appended.

Theoretical work by Kirkwood has determined the course of many experiments in protein chemistry. Continuing his work in the statistical mechanics of liquid solutions, the present report reviews and extends the author's work having to do with the electrical interaction of spherical dipolar ions and ordinary ions and extends the theory to give the limiting law for dipolar ions of elongated shape in solution with common electrolytes. The theory is employed to determine from the solubility data the structure and dipole moments of simple dipolar ions.

One of the most characteristic properties of proteins is denaturation. Although the term is used frequently, its exact meaning is still somewhat elusive. The situation has been greatly improved by recent theoretical work. Representative of it is the present contribution of Eyring and Stearn, in which the process of denaturation is pictured as the simultaneous breaking of a number of weak bonds in the protein molecule with

resulting loss of structure, i.e., a large increase in randomness. In this way solvent is eliminated and the solubility is decreased. Activation in denaturation comes from the breaking of these bridge bonds.

The greatly improved technique of electrophoresis measurements due to Tisclius makes possible many important investigations in this field. Longsworth and MacInnes describe several modifications which they have made in this apparatus and give a summary of the conditions necessary for a disturbance-free process. In an important part of their paper these authors have made a study of some of the boundary anomalies which occur, especially with the concentrated protein solutions with which one has to work in preparative investigations.

With the desire to reduce to the absolute minimum the cost of ultracentrifugal analysis of protein and other systems, McBain is developing new types of "spinning top" ultracentrifuges. These are to be contrasted with the suspended rotor types of air turbine ultracentrifuges which have been developed by Beams, Wyckoff, and Bauer. In the new McBain opaque ultracentrifuges, sedimentation equilibrium and sedimentation velocity may be measured accurately by the application of chemical and biological methods after the ultracentrifugation, thus making optical accessories unnecessary. Methods of immobilization are considered, and typical results are cited.

The isolation of certain of the virus proteins by Stanley several years ago was a remarkable achievement. It appears that at least some members of this group of substances are giant molecules of the nucleoprotein type. As such they have been characterized by physical methods. Lauffer and Stanley outline the results of this nature which have been obtained to date with the tobacco mosaic virus protein, and show how data regarding size and shape may be deduced from them. The results of ultrafiltration, x-ray diffraction, and sedimentation studies are in reasonably good agreement with one another.

As Heidelberger remarks, the inclusion of studies on precipitin and agglutinin reactions in such a symposium as this might have occasioned considerable surprise a few years ago. However, one learns from his report that antigen-antibody reactions may follow quantitative equations which are derived from the mass law, assuming an initial bimolecular reaction followed by competing bimolecular reactions leading to large aggregates. Work of this type has led to the isolation and study of pure antibodies and has provided a new conception of the function of salts in these reactions. The molecular kinetic methods of investigation have proved to be useful in the characterization of the antibodies and in studying the nature of the product formed when toxin and antitoxin react.

The Symposium is concluded with a report by Abramson, Gorin, and

Moyer on the polar groups of protein and amino acid surfaces in liquids. As indicated by electrical mobility and isoelectric point data, changes in the relative strength of acidic or basic radicals of organic ampholytes, such as proteins and amino acids, may or may not occur in surface formation, and such substances are classified on this basis. By means of the interionic attraction theory, calculations are made for the electrical charge of proteins. The importance of these methods of approach to problems in biological systems is described.

The breadth of the subject selected for this Symposium has proved to be a source of considerable difficulty, and it has been found impossible to include detailed discussions of such modes of attack as ultracentrifugal sedimentation and x-ray analysis which, after all, have perhaps given the most direct proofs of the existence of the protein molecule. Also, speculations on the constitution of the protein molecule based upon the amino acid composition are of utmost interest. For instance, Bergmann has postulated a general law expressing the constitution of the proteins as a function of the frequency of the occurrence of the several amino acids. The frequency values obtained from the analytical data and the total number of amino acid residues in the molecule are multiples of the powers two and three. These facts are entirely consistent with the grouping by Svedberg of protein molecules into definite molecular weight classes. Svedberg and Bergmann have presented comprehensive summaries of their work in recent volumes of *Chemical Reviews*. Unfortunately, neither of these scientists was available for this Symposium.

However, it is felt that some success has rewarded our efforts to bring together investigators from representative laboratories, because the summaries produced are uniformly authoritative, at least reasonably comprehensive, and highly suggestive. Some deal with selected physical-chemical aspects of the subject, and others show how this information is being applied in biology, pathology, bacteriology, and other sciences.

We acknowledge with thanks the coöperation of the Editorial Board of *Chemical Reviews* in making it possible to collect the proceedings of this Symposium in a single volume.

PROPERTIES AND STRUCTURE OF PROTEIN MONOLAYERS¹

IRVING LANGMUIR AND VINCENT J. SCHAEFER

Research Laboratory, General Electric Company, Schenectady, New York

Received January 30, 1939

Many proteins, although soluble in water, may be spread on a cleaned water surface to form insoluble monomolecular films. Such films can then be studied in a number of different ways, either while they are still on the water surface or after they have been removed from the water by depositing them on a metal or glass plate.

Some of the properties, such as thickness, reversible compressibility, maximum ease of spreading at the isoelectric point, and insolubility, are quite similar for many proteins of widely different origins, molecular size, and chemical composition. On the other hand, monolayer viscosity and plasticity, anisotropy after shear or linear compression, specific reactivities, expansion patterns, contact angles, and sensitivity to the effect of pH are usually highly specific for particular proteins.

METHODS OF FORMING MONOLAYERS

Proteins may be spread as monolayers in several different ways. When the protein is dissolved in water, a surface monolayer forms spontaneously in less than a second if the solution contains about 0.1 per cent by weight of a protein such as insulin in distilled water at pH 5.8. The film thus formed has a thickness of about 8.5 Å. after 1 min., since the process of spreading at an air-water interface proceeds with enough energy to build up a pressure of 8 dynes per centimeter in 1 min., 10 dynes in 5 min., and 15 dynes in 15 min. Devaux has used this method (5) to study the concentration of a protein solution as a function of the rate of formation of the monolayer.

If the quantity of protein in solution is so small (about 0.1 mg. of protein in a tray with surface area of 200 cm.²) that all of it spread on the water surface would produce five successive monolayers, then, in order to obtain

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

a monolayer of the protein, it is necessary either to wait for a long period of time or to increase the rate of diffusion by thoroughly stirring (14) the substrate. While this technique is quite convenient for reaction studies, the extreme care which must be exercised to prevent surface contamination from diluting the monolayer makes it a less desirable method for general studies of protein monolayers than other methods to be described.

Gorter and Grendel (7) have shown that a dissolved protein may be spread as a monolayer by placing a micropipet in contact with a cleaned water surface and expelling about 50 mm.³ of a 0.1 per cent protein solution from the pipet. If the protein is pure and the water adjusted to the iso-electric point, this amount of protein will produce a monolayer about 500 cm.² in area. We have modified this method by using a nickel sheet about 0.010 in. thick, as long as the width of the tray and as wide as its depth. The protein to be used is spread on this band from a micropipet and the band is lowered at a uniform rate through the cleaned water surface. We believe that this method is particularly efficient when the protein used is of the egg albumin type, which forms a gel-like monolayer. The pressure which is built up during formation of such a monolayer retards the rate of formation somewhat, so that it is quite possible for some of the solution to be dispersed into the substrate. When the protein solution is introduced as a uniform strip across the tray, the spreading proceeds very efficiently on the surface of the upflow of water which follows the lowering of the band.

Still another method of producing good protein monolayers is that developed by Hughes and Rideal (8), who placed a weighed amount of dried protein on the water surface. This method is limited, however, to those proteins which readily spread while in the dried form. Many proteins, such as the globulins, casein, zein, and lactalbumin, to mention a few, spread poorly or not at all when applied in this manner.

Water-soluble proteins produce thick layers, which are visible by interference colors, on the surface of a salt solution (16). A portion of this compact film can then be transferred as a B-layer on a clean nickel plate of known dimensions to another trough containing water at the isolectric point of the protein, and a monolayer can be formed. This method should be particularly useful as a micromethod for purifying and studying proteins which are not available in quantity.

THICKNESS OF PROTEIN MONOLAYERS

Many investigators have shown (6, 19) that the thickness of a protein monolayer, as calculated from the area covered by a weighed amount of protein, is of the order of 10 Å. This calculation can only be approximate, since it assumes that all the protein used has spread to produce the monolayer.

By depositing the protein monolayer on a barium stearate step-plate and measuring its optical thickness (4) it is possible to check the calculated thickness derived from its force-area curve and thus find the efficiency of the spreading technique.

Since different values of pH affect the area of a protein monolayer when identical amounts and methods of spreading are used, changes in observed area as large as ten to one have been interpreted by some investigators as indicating that the protein monolayer becomes thicker as the area decreases. That this is not the case is easily shown by optical measurements of the thickness of deposited monolayers.

Experiments with pepsin (Northrup's crystalline) spread on water having its pH adjusted to a series of different values have given the results shown in table 1. The data of the third column, which give areas per milligram of applied protein (calculated by linear extrapolation to $F = 0$)

TABLE 1
Effect of pH on the area of a pepsin monolayer spread on a water surface

| pH | THICKNESS AT $F = 16.5$ DYNES CM. ⁻¹ | AREA PER MILLIGRAM EXTRAPOLATED TO $F = 0$ | AMOUNT USED | AMOUNT CALCULATED ($d = 1.3$) | RECOVERY |
|-----|-------------------------------------------------------|-----------------------------------------------------|-------------------|---------------------------------------|-----------------|
| | <i>Å.</i> | <i>square meters</i> | <i>milligrams</i> | <i>milligrams</i> | <i>per cent</i> |
| 2.0 | 16. | 0.620 | 0.0416 | 0.0382 | 92 |
| 2.6 | 15. | 0.640 | 0.040 | 0.0376 | 94 |
| 3.0 | 16. | 0.630 | 0.0118 | 0.0110 | 93 |
| 4.2 | 15. | 0.570 | 0.040 | 0.033 | 82 |
| 5.8 | 16. | 0.095 | 0.034 | 0.007 | 20 |
| 7.0 | 15. | 0.001 | 0.200 | 0.003 | 1 |

show, in accord with Gorter's findings, that the maximum spreading occurs near the isoelectric point, pH 2.6. At high pH the area becomes very small. The thickness of the monolayer, as determined after deposition on to a barium stearate plate, remains constant within the accuracy of the measurements.

The total amount of protein in the monolayer (column 5) has been calculated by multiplying the thickness of the dry deposited protein film by its density, assumed to be 1.30, and multiplying this product by the total area of the monolayer on the water. By comparing this with the amount applied to the surface (column 4) we have obtained the percentage recovery as given in the last column.

FORCE-AREA CURVES

The force-area curves for monolayers of proteins obtained by Gorter and others (6, 18, 19) have shown high compressibilities, but, in general,

attention has not been directed to the differences between particular proteins.

Using a surface balance we have determined force-area curves for several proteins as shown in figure 1. The films of insulin, pepsin, and ovalbumin were spread from 0.1 per cent aqueous solutions by our band method, while the wheat gliadin was spread from a 0.1 per cent solution in 70 per cent alcohol using Gorter's method. The pepsin monolayer was

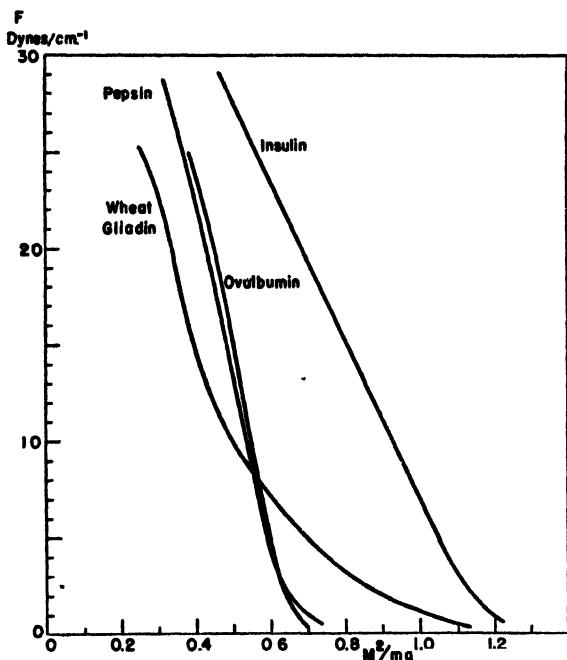


FIG. 1. Typical force-area curves obtained with protein monolayers. Abscissas calculated from measured thickness of deposited monolayers.

formed on water having its pH adjusted to 2.0 by hydrochloric acid, but the others were on distilled water, pH 5.8.

The weight of protein in each monolayer (used in calculating the abscissas in figure 1) was obtained from a measurement of the optical thickness of a pair of layers deposited at $F = 16.5$ dynes per centimeter (dehydrorous AB-film), taking the density to be 1.30. With pepsin, ovalbumin, and insulin the weights of the monolayers found in this way were about 95 per cent of the amounts applied to the surface, but with wheat gliadin only 70 per cent was found on the surface.

The curves for pepsin and ovalbumin resemble one another closely but differ widely from those of insulin and wheat gliadin.

REVERSIBLE COMPRESSIBILITY OF PROTEIN MONOLAYERS

One of the outstanding similarities between protein monolayers spread on a water surface is the high reversible compressibility of the films. Even after being subjected to a momentary decrease in area of 80 per cent, involving a surface pressure of more than 35 dynes per centimeter, an insulin monolayer re-expands to show a complete recovery of its original

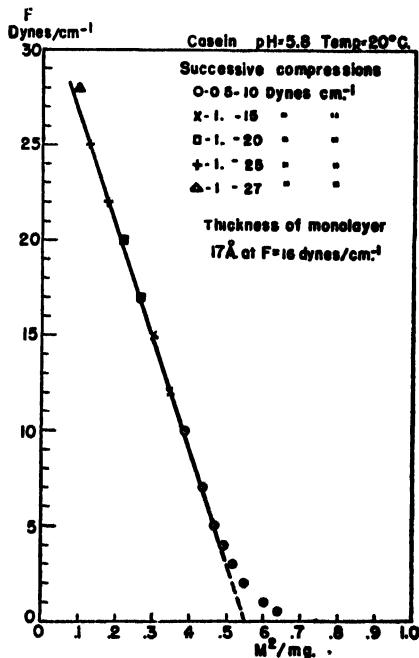


FIG. 2. Reversible compressibility of a casein monolayer. The abscissas represent areas per milligram of protein applied to the surface.

area at $F = 4$ dynes per centimeter. When a film is exposed to this high pressure for several minutes, the recovery is not complete, showing that a permanent change or collapse has occurred.

The force-area curve of a casein monolayer given in figure 2 was obtained from a series of five successive runs in each of which the surface pressure, F , was increased from 1 dyne per centimeter up to a limit indicated in figure 2. The lower portion of each of the curves obtained in this way coincided with the measurements in the previous runs. Thus within the time needed

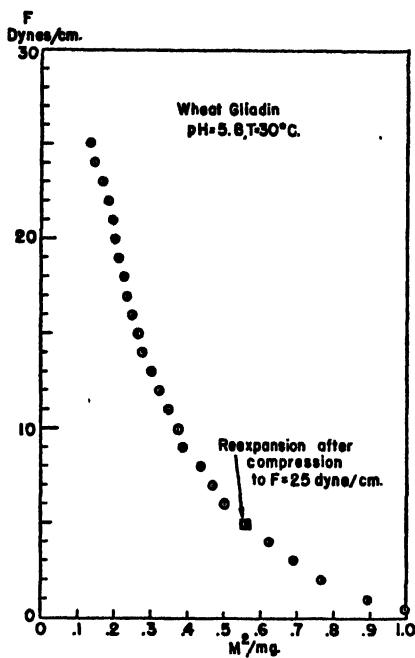


FIG. 3. Reversibility of the compression of a wheat gliadin monolayer between 5 and 25 dynes per centimeter.

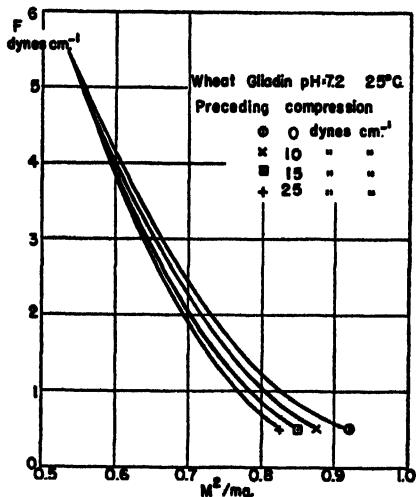


FIG. 4. Irreversible effects observed with a wheat gliadin monolayer at low surface pressures.

for these measurements there was no appreciable collapse or solution of the monolayer.

The abscissas for the curve in figure 2 are expressed in terms of the weights of protein actually applied to the surface. Optical measurements of the thickness of a monolayer deposited at $F = 16$ dynes per centimeter gave $17\text{\AA}.$, which corresponds to 0.45 square meter per milligram; while the curve shown in the figure gives for this value of F an abscissa of 0.28 . Thus to obtain a force-area curve comparable to figure 1 expressed in terms of area per milligram of protein actually present in the monolayer, all the abscissas in figure 2 should be multiplied by 1.60 .

Figures 3 and 4 give the results of similar experiments with monolayers of wheat gliadin. The abscissas represent areas per milligram of applied protein. The points marked by circles in figure 3 were measured at successively increasing values of F up to 25 dynes per centimeter. Upon lowering the pressure to $F = 5$, the film expanded to the same area as was previously observed at this pressure, showing complete reversibility over this range. Some later experiments were made to test an alleged anomalous behavior of gliadin films at very low pressures (18). The force-area curves of figure 4, which represent measurements at intervals of 1 dyne per centimeter, show that the lower part of the curve is distinctly modified by previous compression of the film to pressures even as low as 10 or 15 dynes per centimeter. Since this effect is limited to the portion of the curve below 5 dynes per centimeter, this result is not incompatible with the reversibility indicated by the data of figure 3.

INSOLUBILITY OF PROTEIN MONOLAYERS

It is remarkable that, although many proteins are very soluble in water, the monolayers formed from them are extraordinarily insoluble. Thus a radical change in structure must occur during the spreading. Substances form insoluble monolayers on a water surface only if the surfaces of their molecules have both hydrophilic and hydrophobic parts (9).

Traube (23) showed that the solubilities of adsorbed films of substances containing hydrophobic groups increased by a factor of 3 upon the removal of each successive CH_2 group from the molecule. This was interpreted (9) as an effect of the decrease of surface energy that occurs when CH_2 groups occupy positions in the air-water interface.

The high solubilities in water of such proteins as ovalbumin and pepsin prove that the surfaces of the molecules of these substances contain very few hydrophobic groups. The hydrophobic groups that are responsible for the insolubility of the monolayers must therefore be buried in the interior of the globular molecule of the soluble protein.

The solubility of any adsorbed film, such as a monolayer on a water surface, must theoretically increase if the film is subjected to a surface pressure. The rate of increase can be calculated by the Gibbs equation

$$\frac{dF}{d \ln c} = \sigma kT \quad (1)$$

where c is the concentration of the dissolved substance in the solution, σ represents the number of molecules in the adsorbed film per square centimeter, T is the absolute temperature, and k is the Boltzmann constant, 1.37×10^{-16} ergs per degree.

In the case of condensed films of fatty acids for which $\sigma = 5 \times 10^{14}$ molecules per square centimeter, the solubility should thus increase about fivefold for an increase of 30 dynes per centimeter in the value of F . On the other hand, with a protein of molecular weight of 35,000, the number of molecules per unit area is only about one three-hundredth as great as for the fatty acid, so that there should be a millionfold increase in solubility for each increment of pressure of 1 dyne per centimeter (13a). According to this theory the surface pressure of 25 dynes per centimeter, which was applied to the protein monolayers in the experiments illustrated in figures 2 and 3, should have caused an increase in solubility by a factor of 10^{150} . Actually, however, the complete recovery of the area upon removal of the pressure showed that there was no measurable solubility even at these high pressures. We must therefore conclude that the formation of the monolayer by the unfolding of the globular molecules is an irreversible process to which Gibbs' equation would not apply.

On the basis of studies (10, 11, 12) of the forces that determine solubilities, volatilities, and surface tensions of liquids, it has been calculated (13a) that the presence of hydrophobic groups equivalent to about three hundred CH_2 radicals in each protein molecule would be sufficient to account for the observed insolubility of protein monolayers.

VISCOSITY OF PROTEIN MONOLAYERS

Among the measurable properties of a protein monolayer is its viscosity. This property may be expressed in qualitative terms or in absolute c.g.s. units according to the procedure followed.

A very simple method which serves to classify the monolayer into one of three general classes has been described recently (20). It involves the formation of the protein monolayer, followed by the expansion of its central region with a spreading oil which reveals the degree of cohesion by which the units in the monolayer are joined to each other. The greatest difference found among monolayers of different proteins occurs when they are

under low surface compression of the order of 0.5 to 1 dyne per centimeter. The three general patterns observed have been termed star-like, rough circular, and smooth circular. These types are illustrated in figure 5.

The star-like pattern expands to form geometrical figures, sometimes as



FIG. 5 General types of expansion patterns: a, star-like (egg albumin); b, rough circular (papain), c, smooth circular (insulin)

TABLE 2

Absolute viscosities of protein monolayers

$M = 131.8$, $D = 4.5$, $\text{pH} = 5.8$, $T = 25^\circ\text{C}$; $t = \text{about } 34 \text{ sec}$

| PROTEIN | μ_v IN GRAMS PER SECOND | | | | | |
|-----------------------|-----------------------------|------------------|-------------------|---------------------|-------------------|---------------------|
| | $F = 2$ dynes | $F = 6$ dynes | $F = 10$ dynes | $F = 16.5$ dynes | $F = 19$ dynes | $F = 29.5$ dynes |
| Casein | 0.0036 | 0.010 | 0.015 | 0.4 | 12 | |
| Edestin | 5 | 26 | 52 | | | |
| Egg albumin | 0.15 | 0.28 | | | | |
| Gliadin* | 0.001 | 0.003 | 0.01 | 0.8 | | |
| Gliadin acetate | 0.001 | 0.007 | 0.016 | | | |
| Hemoglobin | 0.024 | 0.12 | 0.37 | | | |
| Horse globulin | 120 | 210 | 340 | | | |
| Insulin | 0.004 | 0.028 | 0.10 | | | |
| Papain | 0.08 | 0.18 | 0.38 | | | |
| Pepsin | 0.68 | 0.75 | 1.5 | 150 | | |
| Pepsinogen | 0.110 | 0.17 | 0.46 | | | |
| Tobacco seed globulin | 0.048 | 0.20 | 0.36 | 140 | | |
| Trypsin | 0.009 | 0.23 | 0.97 | | | |
| Trypsinogen | 0.40 | 1.0 | 5.5 | | | |
| Zein | 0.003 | 0.003 | 0.003 | | | |

* At pH = 7.2.

a conventional five-pointed star but at times with points ranging in number from three to six or more. The star-like form at $F = 2$ dynes ranges in viscosity from 0.024 to 120 c.g.s. units. Proteins having this form are egg albumin, pepsin, or tobacco seed globulin.

The rough circular pattern shows a circular internal outline which upon

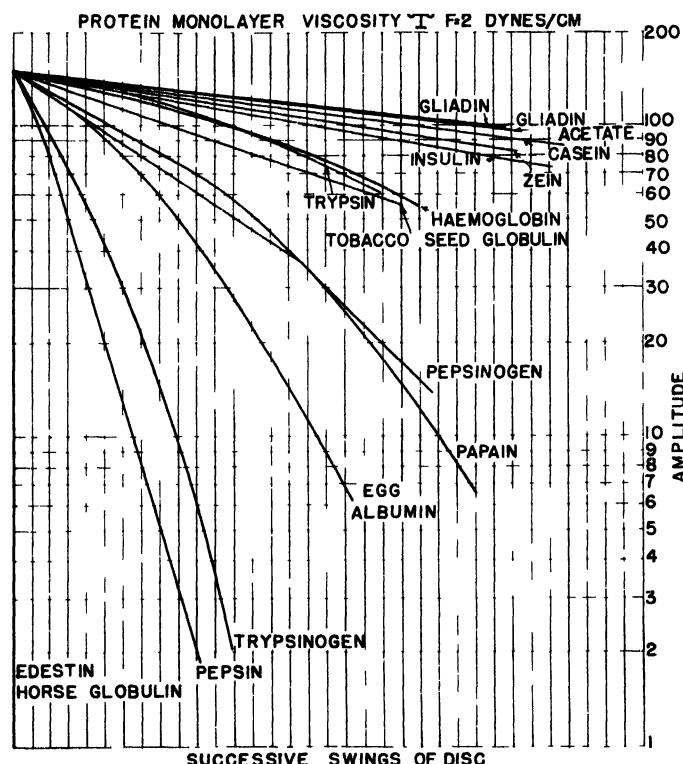


FIG. 6 Viscosities of protein monolayers $F = 2$ dynes per centimeter Oscillation method $\text{pH} = 5.8, T = 22^\circ\text{C}, M = 131.8, D = 4.5, t = \text{about } 31 \text{ sec}$

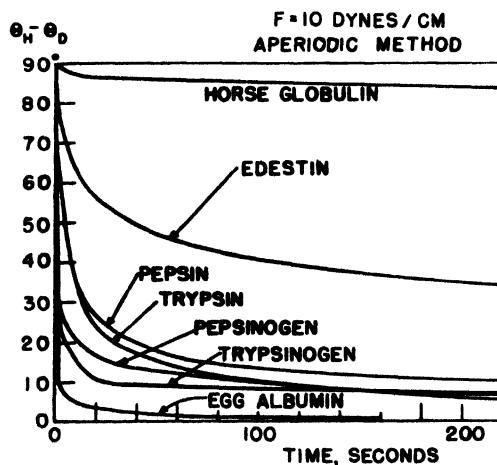


FIG. 7 Viscosities of protein monolayers $F = 10$ dynes per centimeter Aperiodic method $\text{pH} = 5.8, T = 22^\circ\text{C}, D = 4.5$

close examination discloses an irregular or roughened edge at the boundary of the oil and the protein. The viscosity range of this form is from 0.009 to 0.081 c.g.s. units at $F = 2$ dynes for those investigated. Typical proteins of this type are trypsin, papain, and wheat gliadin (when the latter is spread in its dry form).

The smooth circular type shows a circular internal pattern with the oil-protein boundary perfectly regular and smooth. Such patterns are found with insulin, casein, and zein, and denatured proteins and have a viscosity range at $F = 2$ dynes of 0.001 to 0.004 c.g.s. units.

The viscosity of monolayers of all types can be determined quantitatively by a relatively simple method previously described (14). Table 2 gives a summary of viscosities of a number of proteins measured under definite degrees of compression. These data are presented primarily to indicate the interesting ranges existing in protein monolayers. More detailed work is planned to investigate the importance of pH, temperature, method of spreading monolayers, and the effect of various modifications to the native protein which occur when it is subjected to heat, acid, ultraviolet irradiation, shaking, and other methods of "denaturation". We have shown qualitatively (20) that a large change in viscosity occurs when a protein such as pepsin is subjected to any of these agents.

In general, proteins producing monolayers having smooth and rough circular expansion patterns may be investigated by the oscillation method. This involves the measurement of the decrease in amplitude of the successive swings of a disk lying in the surface of the water in contact with the monolayer and suspended at the end of a calibrated fiber. The disk is provided with a bracket upon which is placed a rod having a known moment of inertia. The fiber is given an initial displacement by the torsion head, and the decrease in amplitude of the disk is noted. Figure 6 shows typical curves obtained by this method.

While the viscosity of a few of the proteins showing the star-like expansion pattern may be measured by the method just referred to, most of them have such high viscosities as to necessitate the use of the aperiodic method. This uses the same viscosimeter but involves the comparison of the relative movement of the suspended disk in contact with the monolayer with the initial displacement of the torsion head as a function of elapsed time. In this method the rod used to give a known moment of inertia to the system is removed. Figure 7 shows a series of curves obtained in this way at $F = 10$ dynes per centimeter.

THE VISCOSITIES OF MONOLAYERS OF LIPO-PROTEIN COMBINATIONS

Schulman and Rideal (21) have described experiments which show a very interesting interaction between a protein and a lipoid substance.

Using a mixture of cholesterol and wheat gliadin on water at pH 7.2 in a ratio of one part of cholesterol to four parts of protein, they demonstrated that at a critical pressure the protein was apparently squeezed out of the monolayer on the surface into the substrate and that certain properties of the surface layer were altered if this critical pressure were exceeded.

A wheat gliadin monolayer at low pressures is liquid, but as it is compressed it changes from a liquid to a gel-like solid. Cholesterol, on the other hand, is a liquid film under all compressions up to 30 dynes per centimeter. Schulman and Rideal showed that as the complex film of lipo-protein was compressed it resembled the protein monolayer, but at a critical pressure of about 22 dynes per centimeter the film suddenly liquefied and on further compression resembled the cholesterol film. This

TABLE 3

*Absolute viscosities of protein, lipo-protein, and lipoid films
 $M = 131.8$; $D = 4.5$; pH = 7.2; $T = 25^\circ\text{C}.$; $t = \text{about } 34 \text{ sec.}$*

| CHOLESTEROL parts by weight | GLIADIN parts by weight | μ_s IN GRAMS PER SECOND | | | | | | | |
|--------------------------------|----------------------------|-----------------------------|----------|------------|----------|------------|------------|------------|------------|
| | | $F = 2$ | $F = 10$ | $F = 16.5$ | $F = 19$ | $F = 20.5$ | $F = 16.5$ | $F = 17.5$ | $F = 20.5$ |
| 1 | 0 | 0.002* | 0.002 | 0.019 | 0.002 | 0.002 | | | |
| 1 | 2 | 0.002 | 0.02 | 0.04 | 0.04 | 0.002 | | | |
| 1 | 4 | 0.001 | 0.001 | 0.76 | 0.16 | 0.05 | | 0.49 | 0.04 |
| 1 | 8 | 0.002 | 0.002 | 0.36 | 0.06 | 0.02 | 0.36 | | |
| 1 | 16 | 0.006 | 0.006 | 0.18 | 0.85 | 1.25 | | | |
| 0 | 1 | 0.001 | 0.01 | 0.83 | | 54. | | | |

Measurements of viscosities were made in successive order as shown.

* The low values of μ_s are uncertain within about ± 0.005 .

sequence was shown to be reversible; if the surface pressure were lowered, the liquid film showed a re-gelation at the same critical pressure.

We have repeated these experiments with rough quantitative determinations of viscosity, varying the ratio of gliadin to cholesterol over a wide range, and have thus obtained the data given in table 3.

A monolayer of pure cholesterol, which is relatively incompressible, has a low viscosity even at a high surface pressure, but the viscosity of a gliadin monolayer undergoes a very rapid increase at pressures above 10 dynes per centimeter. The viscosities of the mixed films rise to a broad maximum between 15 and 20 dynes per centimeter and fall to low values at $F = 30$. This effect is most marked with the 1:4 mixture, but even as little as one part of cholesterol to sixteen parts of gliadin produces a very

great decrease in viscosity at high pressures, the cholesterol apparently acting as a lubricant within the gliadin monolayer.

ANISOTROPY IN PROTEIN MONOLAYERS

When a monolayer of a protein such as egg albumin, urease, casein or edestin is compressed for a few minutes by a motion of a barrier which raises the pressure to about 35 dynes per centimeter, so that some collapse occurs, a permanent anisotropic condition is produced which can be demonstrated in several ways.

Thus if the pressure is lowered to about 20 dynes per centimeter, the monolayer can be split into parallel sheets by using a concentrated oxidized oil. These molecular strips can be lifted from the surface of the water by placing a glass rod or similar article underneath one end and slowly raising it through the water surface. The strips come off as fibers which are almost invisible. Fibers can also be removed by simply inserting the rod into a compressed film of these proteins and slowly raising it through the surface. When this is done, the fiber always comes off parallel to the barrier which has compressed the monolayer. Not all protein monolayers under compression show this property. Insulin, for example, when cracked with indicator oil breaks into small segments which will not produce fibers.

The fact that the insulin expansion pattern is reversible and shows a smooth circular pattern at $F = 2$, a star-like pattern at $F = 25$, and a smooth circular pattern again when the pressure is lowered suggests that its structure differs in important features from that of the fiber-forming proteins. Casein, for example, which is characterized by a circular pattern and a viscosity lower than that of insulin at $F = 2$, has a fibrous structure at $F = 30$. When the pressure is decreased, the expansion pattern shows that the casein now has a coherent gel-like structure, even at $F = 2$.

By cutting a portion of the spread monolayer into various shapes, such as a square, rectangle, circle, etc., we can study the effect of the presence or absence of anisotropy under compression. The fragments are made with a wire about 0.015 in. in diameter which, after being cleaned in a Bunsen flame, is coated with indicator oil. The wire is lowered through the monolayer, and the compression is decreased. As the oil spreads, the monolayer is sharply outlined. Any inequality produced by the initial compression is immediately indicated by a distortion of the original configuration. These experiments are best performed with a suitably illuminated trough having a black bottom (20).

A protein monolayer on water when compressed equally in all directions by an indicator oil or piston oil shrinks uniformly toward its geometrical center and is isotropic. Thus, the piston oil applied at one end of the monolayer is found not only at this end but also along both sides of the

tray. Because of this phenomenon, we make a practice of applying compression by placing the piston or indicator oil in one corner of the tray. This produces a uniformly compressed, rectangular monolayer having oil on only two sides.

INDICATOR AND PISTON OILS USED FOR STUDYING MONOLAYERS

We have often referred to indicator and piston oils as convenient tools in the study of monolayers. Dr. K. B. Blodgett (1) has described the method for preparing and using highly oxidized oil. For most experiments it is quite practical to use old automobile crank-case oil. Fresh oil is not satisfactory. The sludge in the oil may be removed by filtering or settling. The oxidized oil when diluted with a non-spreading oil such as Petrolatum or Nujol will produce stable films of uniform thickness, which in certain mixtures will show interference colors.

A sample of used oil was placed on a cleaned water surface and spread at $F = 0$ to produce a film which, viewed at approximately 45° , showed a faint silver-grey color. The addition of an equal amount of Petrolatum produced a film having a faint yellow color of the first order. When two parts of Petrolatum to one of oxidized oil were blended, the resulting color was a dark yellow of the first order; three parts of Petrolatum and one of oxidized oil produced a first-order blue. Four parts of Petrolatum to one of oxidized oil resulted in a yellow again, and with a five-to-one ratio red was produced, these colors being of the second order.

When an indicator oil, which will spread on clean water if unrestricted to give a dark yellow, first-order color, is limited in area, the color produced by the increase in thickness will be related to the decrease in area. In order to decrease the area, it is necessary to exert a definite amount of pressure on the oil film. If a trough equipped with a surface balance is used, it is quite easy to calibrate a given oil sample in terms of color observed for pressure exerted on the floating barrier. This has been done with various samples of oils which we use to illustrate the possible ranges of pressure. These are given in table 4. Although such oils are fairly stable if protected from light, they should be checked occasionally for changes which might occur.

With a series of oils such as those indicated in table 4, we have a very convenient and simple method for producing compressions up to 30 dynes per centimeter upon a monolayer spread on water.

To complete the series a few piston oils, such as those given in table 5, may be used. These are purified by shaking a good grade of commercial oil with activated fuller's earth or similar substance at about 100°C . The oils may then be clarified by centrifuging or allowing the fuller's earth to settle. These oils have the property of spreading as liquid monolayers,

the excess oil forming lenses which serve as reservoirs and produce a constant, known pressure upon anything floating on the water surface.

Indicator oil used in conjunction with piston oil can be used in many ways other than those suggested.

TABLE 4

The relation of the color of oxidized oil to the pressure exerted against a monolayer on water

| COLOR $i = 45^\circ$ | <i>F IN DYNES PER CENTIMETER</i> | | | | | |
|-------------------------|----------------------------------|-----------------------------------------------------|------|------|-----|--------------------|
| | Oxidized oil | Oxidized lubricating oil diluted with petrolatum | | | | Crank- case oil |
| Invisible | 0-5 | | | | | |
| Silver-grey | 11. | 0. | | | | 2-5 |
| Light yellow | 14. | 3. | | | | 7. |
| Dark yellow | 15. | 5. | 0. | | | 8.5 |
| Red-brown | 15.5 | 7. | 2.5 | | | 9.5 |
| Blue | 16.0 | 8. | 4. | 0. | | 10.5 |
| Blue-green | | 9. | 5. | 2. | 0. | 11.2 |
| Yellow | 17.0 | 9.5 | 6. | 3.5 | 0. | 11.8 |
| Red | 18. | 10.5 | 7. | 5. | 1. | 12.5 |
| Blue | 18. | | 8. | 6. | 2. | 13.0 |
| Green | 18. | 11. | 8.2 | 6.5 | 3.3 | 13.5 |
| Red | 18. | 12.5 | 9.5 | 7. | 5. | 14.5 |
| Green | | 13.5 | 10.0 | 8.5 | 6.0 | 15.0 |
| Red | | 13.7 | 10.5 | 9.0 | 7.0 | 15.5 |
| Green | | | 11.2 | 9.5 | | 16. |
| Visible lens | 20-30 | 20-30 | 15. | 13.5 | 12. | 20-30 |

TABLE 5

Force exerted by piston oils spread on a water surface

| PISTON OIL | <i>F IN DYNES PER CENTIMETER</i> |
|--------------------------|--------------------------------------|
| Tricresyl phosphate..... | 9.5 |
| Rape seed oil..... | 10.5 |
| Castor oil..... | 16.5 |
| Neatsfoot oil..... | 19.0 |
| Oleic acid..... | 29.5 |

DEPOSITION OF MONOLAYERS

In determining the thickness of protein monolayers several precautions must be observed. Since most proteins undergoing compression develop considerable rigidity, the monolayer does not flow at a uniform rate toward the plate upon which it is to be deposited. This difficulty can be overcome by using a plate of the same width as the monolayer. Since it is not

convenient to produce a step-plate of this size, we use chromium-plated steel plates of two sizes. One plate is made as wide as the tray (allowing 1-mm. clearance on either side), and the other plate is of a convenient size (5 cm. x 2 cm. x 1 mm.). The former is held by a small permanent magnet of sufficient strength to cause the prepared plate to adhere firmly to the side of the larger plate. Glass or a non-magnetic metal could also be used for the larger plate.

We prefer to produce the initial compression by using a calibrated indicator oil. This spreads and outlines the boundary of the protein monolayer; it also prevents the invisible piston oil from contaminating the monolayer. The indicator oil is then compressed to a force approaching that of the piston oil to be used. The piston oil is added and drives the protein monolayer on to the prepared plate (2, 3, 4), where it is deposited as a hydrous AB-film (17). When monolayers are to be deposited at pressures higher than $F = 16$ dynes, concentrated oxidized lubricating oil may be used as the indicator oil. This, no matter how highly compressed, can be seen either as a colored film or as a line of visible oil.

A monolayer in contact with a film spread from a specific sample of calibrated indicator oil will be under a definite compression directly related to the color of the oil film (see table 4). If the color is held constant, the monolayer may be deposited as a hydrous AB-film. To do this, the synchronous movement of the surface barrier compressing the film and the dipping plate is necessary.

A monolayer may also be deposited by lowering the prepared plate face downward upon the top surface of the protein monolayer. A film so deposited has been termed a lifted film, symbolized by A_L . Best results are obtained if one corner of the prepared plate, held almost parallel to and about 1 mm. from the film surface, is touched to the monolayer. In this way the formation of an air bubble between the prepared plate and the monolayer is prevented. After the monolayer is deposited, the surrounding film is scraped away, and the face of the plate, now bearing the deposited A -layer, is raised through the cleaned surface. Care must be exercised that the edges of the plate are free from surface-active materials. This is easily checked by noting any change in color of the indicator oil before and after the film is deposited. The deposition of an A_L -film is particularly applicable to the deposition of films under low compression, since no surface movement of the monolayer is necessary.

THE SURFACES OF PROTEIN MONOLAYERS

In a previous paper (17) data were given which showed certain relationships found when successive protein monolayers were deposited on a prepared plate. Using our improved optical method of measuring incre-

ments of thickness, we have repeated these experiments with monolayers of insulin (Lilly).

PRAAA-layers were deposited by spreading a monolayer on the water surface, applying pressure, and then lowering a prepared step-plate bearing a multilayer of barium stearate into the water. The protein monolayer was deposited on the down trip. Before the plate bearing the deposited monolayer was withdrawn through the cleaned water surface, the remaining film on the water was scraped off. This process was repeated for each succeeding layer added.

Successive PRBBB-layers are deposited by a reversal of the above-described process. The plate is lowered through a cleaned water surface, a protein monolayer spread on the surface, compressed, and deposited as the plate is withdrawn from the water. PRABAB-layers are deposited in a manner similar to that used in building ordinary barium stearate multilayers. However, the water between the two layers must be allowed to evaporate before the next pair is deposited, otherwise the B-layer returns to the water surface.

We first thought it necessary to condition the monolayer with divalent salts in order to prevent the return of the B-layer to the water as successive AB-layers were deposited. We have since found that protein monolayers can be built without this conditioning step by increasing the speed of the down trip to about 30 cm. per second.

The observed thicknesses show that with a surface pressure of $F = 16.5$ dynes per centimeter the PRABAB-layers increased 22 \AA . each trip, an average increment of 11 \AA . per layer. The PRAA combination showed an increase of 10 \AA . per layer for the first three trips, with no further increase thereafter. Each layer of the PRBB combination gained 10 \AA . for each trip made; there appeared no limit in this case to the number which could be deposited.

When a multilayer of insulin consisting of successive dehydrous AB-layers is slowly lowered into the water, the outer B-layer returns to the water surface. If this layer is scraped off, the plate withdrawn and dried, the multilayer shows the loss of a single layer. If this process is repeated the second time, the multilayer when dried shows the loss of two additional layers. This may be continued until all the pairs of BA-layers are stripped off successively down to the PRA protein monolayer. This layer cannot be removed in the manner described. Each of these BA-layers consists of two monolayers which are held together by contact between their hydrophobic surfaces. On the other hand a pair of monolayers that form an AB-layer are joined by their hydrophilic surfaces. Apparently the penetration of water between the hydrophilic surfaces within the AB-layers accounts for the manner in which the multilayer was stripped from

the plate as BA-layers and not in the AB combination in which they were applied.

"Breath figures" produced when moist air is brought into contact with a slightly chilled slide bearing deposited monolayers show certain striking differences between protein layers deposited in various ways.

If we breathe upon a chilled chromium-plated surface, which has previously been polished with Shamva and is therefore hydrophilic, we find that moisture condenses as a fine fog-like deposit consisting of minute drops which scatter light strongly but give no interference colors. A quite similar fog-like deposit is produced when one breathes upon deposited films of insulin of the following types: PRA, PRB, PRAAA, and PRABABA. It appears, therefore, that these films act as if the outside surface were hydrophilic.

From the orientation of the hydrophobic and hydrophilic groups in the protein monolayer on the water surface before deposition, one would expect an A-layer to be exotropic (13), that is, it should have its hydrophilic groups turned outward (away from the plate). Similar considerations would indicate that B-layers should be endotropic, having their hydrophilic groups turned inwards (toward the plate).

The fact that the PRB-films are found to act as though they have hydrophilic surfaces is presumably accounted for (13) by the overturning of the layer which occurs because the underlying hydrophobic R-layer tends to anchor the hydrophobic parts of the B-layer and so makes it easy for the hydrophilic parts to come into contact with water when this is condensed on the surface.

Films of the type PRABAB and PRBBB, when breathed upon, give an entirely different kind of breath figure. At first sight there appears to be no breath figure at all, since there is no fog-like deposit which scatters light. Closer examination, however, shows that the surface gives bright interference colors which indicate the presence of a uniform film of water of thickness even greater than the wave length of light. Evidently the water condenses as a thick film under the B-layer that forms the surface of the film, the phenomenon being very similar to that observed when the outside B-layer is stripped off of a PRABAB-film by dipping this slowly into water. The formation of these thick uniform condensates is therefore not incompatible with an endotropic orientation of the B-layer.

These experiments seem to indicate definitely that B-layers and A-layers normally retain the dorsiventral character of the protein monolayer on water, so that the outer surface of an A-layer tends to be hydrophilic while that of a B-layer is hydrophobic. In the case of the PRB-films,

however, the phenomenon is complicated by the overturning which may occur if the films are not suitably anchored by the underlying layers.

REACTIVITY OF PROTEIN MONOLAYERS

A previous paper (15) has discussed the reactivity of certain proteins when spread as monolayers. We concluded that in the case of pepsin the process of spreading the protein as a monolayer and depositing it on a plate did not seriously interfere with its ability to clot milk. We found, how-

TABLE 6
Activities of urease monolayers
T = 25°C.; citrate buffer; pH = 6.5

| BASE | TYPE | <i>F</i> | THICKNESS | UREASE | UREASE UNITS PER SQUARE CENTI- METER | UREASE UNITS PER GRAM IN SOLUTION |
|------------------------------|------|----------|-----------|--------------------|--------------------------------------------------|--------------------------------------------|
| | | | | UNITS† PER GRAM | | |
| PR. | AL | 1. | 16 | 11,100 | 0.0023 | 0 |
| PR. | AL | 10.5 | 17 | 9,100 | 0.0020 | 800 |
| PR. | AL | 16.5 | 31 | 2,400 | 0.0010 | 0 |
| PR. | AL | 29.5 | 49 | 2,700 | 0.0016 | |
| PR. | AL | 1. | 22 | 10,100 | 0.0023 | |
| PR. | AL | 1. | 21 | 7,330 | 0.0016 | 800 |
| PRC (thorium desoxycholate). | AL | 1. | 18 | 5,200 | 0.0013 | 0 |
| <hr/> | | | | | | |
| PRC (thorium desoxycholate). | AL | 29 | 21* | 16,800 | 0.0041 | 0 |
| (same film, second dip) | AL | 29 | | 16,800 | 0.0041 | 16,800 |
| PRC (ThO ₄) | S | | 15 | 2,000 | 0.0004 | |

* This observed thickness is 28 Å. less than that observed on a PR barium stearate film.

† A urease unit is defined as that amount of urease which will form from urea phosphate 1 mg. of nitrogen at 20°C. and at pH 7 in 5 min.

Nitrogen was determined with Nessler's reagent.

ever, that the monolayer left the plate, being removed presumably by something in the milk, since washing with distilled water or a buffered solution failed to remove it. A resynthesis of the pepsin, as the protein leaves the slide, is suggested as having occurred.

A few preliminary qualitative results were also announced concerning the enzyme urease. Further quantitative work (13a) has indicated that although the monolayer shows the ability to convert urea into ammonia it has an efficiency in urease units (22) of only 5 per cent when compared to that of a similar amount added in bulk to a urea solution. Examination

of the monolayer shows that considerably more light is scattered from the surface of a deposited urease monolayer than from one of insulin or pepsin applied under the same conditions. Thus it is possible that the observed activity is due primarily to entrapped globular molecules of the urease which have not been broken down into a monolayer. This possibility is further strengthened by the fact that a monolayer left on water for 15 min. (a period which should produce greater spreading if there were entrapped unspread globular molecules) shows less than 2 per cent activity. Table 6 contains some of the quantitative data obtained.

It will be noted that as a result of the quantitative measurements which we have made with the urease monolayers our previous conclusions must be modified, particularly those concerning the effect of conditioned surfaces on the activity of urease. These data indicate that the monolayer is less firmly held by the thorium desoxycholate surface than by the barium stearate surface and therefore becomes detached from the slide and diffuses into the urea solution. This is shown by the observed decrease in thickness of the monolayer as well as by the continuation of urea conversion after the monolayer is removed. Our previous technique would not have disclosed this condition.

Further work is planned with particular attention to the amount of light scattered by single monolayers and the loss of activity with aging of the monolayer on the water surface.

When a monolayer of catalase is spread on a water surface, transferred to a plate and immersed into a solution of hydrogen peroxide, oxygen accumulates on the surface of the monolayer as visible bubbles of gas 1 mm. or more in diameter. When the slide is withdrawn from the solution, both the bubbles and the catalase monolayer beneath each bubble disappear.

A monolayer of catalase deposited as an A_L -layer at $F = 1$ dyne per centimeter was quite hydrophilic and after a thorough washing was found to be about 23 Å. thick. After rewetting, by dipping through a cleaned water surface, the plate bearing the monolayer was dipped into a 3 per cent hydrogen peroxide solution. After the bubbles of oxygen had become about 1 mm. in diameter the plate was agitated to remove the bubbles and permit new ones to form. After this was done four times the bubbles stopped forming. The plate after removal from the solution was found to be relatively hydrophobic, with the bubble locations visible as regions of decreased thickness.

When a catalase monolayer was deposited on top of a barium stearate step-plate conditioned with thorium desoxycholate and treated in the manner just described, an interesting difference was observed. As the bubbles started forming, they left the surface of the monolayer before reaching

one-quarter the size of those formed on a catalase monolayer deposited on barium stearate. As far as could be determined, the total activity of the film before the evolution of oxygen stopped (a period of about 15 min.) was of the same order of magnitude,—roughly 700 cc. of oxygen per milligram or 0.2 cc. of oxygen per square centimeter of monolayer.

Further work is planned to determine whether in the case of catalase as with urease the activity of the monolayer is possibly due to unspread molecules enmeshed in the fabric of the spread monolayer.

STRUCTURE OF PROTEIN MONOLAYERS

The solubility and other properties of the globular proteins, together with the insolubility and irreversible formation of the monolayers, indicate that the globular proteins have a symmetrical, highly organized cage-like structure which breaks down or unfolds during this spreading process, permitting the hydrophobic groups to come into contact with the air-water interface.

A study of the properties of several long-chain polymers, such as methyl methacrylate and polyvinyl acetate, has shown that the force-area curves, the expansion patterns, and the viscosities of these monolayers resemble closely those given by proteins. The monolayers of some of these polymers when compressed become anisotropic and yield fibers much like those observed with egg albumin.

Protein monolayers thus consist mainly of polypeptide chains which are anchored to the air-water interface at intervals along their length by hydrophobic groups contained in the side chains (13a). The structure is like that of a net made to float on the surface of water by corks distributed over the surface of the net. The hydrophilic parts of the polypeptide chains remain in contact with and are surrounded by water but are not free to go far from the surface because they form parts of chains attached to the surface. The hydrophobic groups by their Brownian movement tend to distribute themselves over the surface like the molecules of a two-dimensional gas, but this expansion is constrained by the elastic properties of the long chains which result from the flexibility of the chain (17a).

The marked insolubility of the protein monolayers even under high compression results from the great length of the polypeptide chains in the monolayers. The incomplete reversibility shown in the force-area curve of a wheat gliadin monolayer at low compressions, figure 4, may be explained by the presence of a small proportion of relatively short-chain degradation products produced during the spreading of the protein, which are driven into solution by subjecting the film to a high pressure. We plan to study this phenomenon with films given by more strongly denatured proteins.

The dorsiventral properties of protein monolayers shown in our studies of deposited films indicate that the deposited monolayers resemble a fabric. Cross-linkages which give to the monolayer this fabric-like character may be due to the adherence between hydrophobic groups or may involve the cyclol bonds which have been postulated in the structure of the globular protein (25).

Quantitative studies of viscosities, compressibilities, elasticities, etc. of protein monolayers should help greatly to throw more light not only on the structure of monolayers themselves but also on that of the globular proteins from which they are derived.

REFERENCES

- (1) BLODGETT, K. B.: J. Optical Soc. Am. **24**, 313 (1934).
- (2) BLODGETT, K. B.: J. Am. Chem. Soc. **57**, 1007 (1935).
- (3) BLODGETT, K. B.: J. Phys. Chem. **41**, 975 (1937).
- (4) BLODGETT, K. B., AND LANGMUIR, I.: Phys. Rev. **51**, 964 (1937).
- (5) DEVAUX, H.: Compt. rend. **200**, 1560 (1935).
- (6) GORTER, E.: Proc. Acad. Sci. Amsterdam **37**, 788 (1934).
- (7) GORTER, E., AND GRENDEL, F.: Trans. Faraday Soc. **22**, 477 (1926).
- (8) HUGHES, E., AND RIDEAL, E. K.: Proc. Roy. Soc. (London) A**137**, 62 (1932).
- (9) LANGMUIR, I.: J. Am. Chem. Soc. **39**, 1848 (1917).
- (10) LANGMUIR, I.: Colloid Symposium Monograph **3**, 48 (1925). The Chemical Catalog Co., Inc., New York (1925).
- (11) LANGMUIR, I.: Colloid Chemistry, edited by J. Alexander, p. 536. The Chemical Catalog Co., Inc., New York (1926).
- (12) LANGMUIR, I.: Chem. Rev. **6**, 451 (1929).
- (13) LANGMUIR, I.: Science **87**, 493 (1938).
- (13a) LANGMUIR, I.: Cold Spring Harbor Symposia Quant. Biol. **6**, 171 (1938); see also pages 136 and 186.
- (14) LANGMUIR, I., AND SCHAEFER, V. J.: J. Am. Chem. Soc. **59**, 2400 (1937).
- (15) LANGMUIR, I., AND SCHAEFER, V. J.: J. Am. Chem. Soc. **60**, 1351 (1938).
- (16) LANGMUIR, I., AND SCHAEFER, V. J.: J. Am. Chem. Soc. **60**, 2803 (1938).
- (17) LANGMUIR, I., SCHAEFER, V. J., AND WRINCH, D. W.: Science **85**, 76 (1937).
- (17a) MARK, H.: Nature **141**, 670 (1938).
- (18) MITCHELL, J. S.: Trans. Faraday Soc. **33**, 1129 (1937).
- (19) PHILIPPI, G. T.: On the Nature of Proteins. Noord Hollandsche Uitgeversmaatschappij, Amsterdam, Holland (1936).
- (20) SCHAEFER, V. J.: J. Phys. Chem. **42**, 1089 (1938).
- (21) SCHULMAN, J. H., AND RIDEAL, E. K.: Proc. Roy. Soc. (London) B**122**, 46 (1937).
- (22) SUMNER, J. B.: J. Biol. Chem. **69**, 435 (1926).
- (23) TRAUBE, J.: Ann. Physik **265**, 27 (1891).
- (24) WRINCH, D. M.: Proc. Roy. Soc. (London) A**161**, 505 (1937).
- (25) WRINCH, D. M.: Phil. Mag. **25**, 705 (1938).
- (26) WRINCH, D. M., AND LANGMUIR, I.: J. Am. Chem. Soc. **60**, 2247 (1938).

SOME PHYSICAL-CHEMICAL CHARACTERISTICS OF PROTEIN MOLECULES^{1, 2}

EDWIN J. COHN

Department of Physical Chemistry, Harvard Medical School, Boston, Massachusetts

Received January 30, 1939

I. INTRODUCTION

The importance of developing an adequate chemistry of the proteins has, perhaps, never been doubted. These vast molecules could not, however, be adequately described by the methods available to earlier generations. The analytical chemistry of the early part of the 19th century accurately revealed, it is true, their elementary composition. A century ago Liebig and Mulder knew that proteins were rich in nitrogen, and contained, besides carbon, hydrogen, and oxygen, small amounts of sulfur and of phosphorus. A few characteristic analyses of this period are worth republishing (see table 1).

The elementary composition of most proteins was thus early estimated to be so nearly alike as not to suggest the vast chemical and biological differences that can be ascribed to the different molecules of this group. Only the analyses for sulfur and phosphorus showed wide variations from protein to protein and led to the controversy as to whether these elements were indeed part of the protein molecule. No sulfur-containing amino acid had been isolated from a protein hydrolysate³ in 1838, but Mulder

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

² The points of view developed in this communication have in part been derived from the discussions held in our seminar. I am therefore indebted to George Scat- chard, John G. Kirkwood, and John D. Ferry for considering with us intermolecular forces; to Hans Mueller, Jeffries Wyman, Jr., J. L. Oncley, John D. Ferry, and J. Shack for considering dielectric constant measurements and their interpretation; to John T. Edsall and John W. Mehl for considering diffusion, viscosity, and double refraction of flow; to Ronald M. Ferry and A. A. Green for considering electromotive force and electrophoretic mobility; to Norval F. Burk for considering osmotic pressure; to J. L. Oncley and Jacinto Steinhardt for considering sedimentation methods of determining molecular weight; and to T. L. McMeekin and Jesse P. Greenstein for considering the amino acid composition of the proteins.

³ Wollaston (158) discovered cystine in 1810, but it was not isolated from a protein until eighty years later (81, 97).

calculated that on the basis of the analyses for sulfur and phosphorus in table 1, and the atomic weights in use at the time, that the molecular weights of these proteins would be over 50,000.

The isolation of the amino acids from the hydrolysates of proteins occupied the attention of the great organic chemists of the 19th and the early 20th centuries. Something over twenty amino acids are now recog-

TABLE 1
Elementary composition of proteins

| COMPOSITION | | | | | | | PROTEIN | FORMULA | | | | | | MOLECULAR WEIGHT |
|-------------|---|---|---|---|----|---|---------|---------|---|---|---|---|----|------------------|
| C | H | N | O | S | Fe | P | | C | H | N | O | S | Fe | P |

According to Mulder in 1838* (99)

| | | | | | | | | | | | | | | | |
|-------|------|-------|-------|------|--|------|---------------|-----|-----|-----|-----|---|--|---|--------|
| 54.56 | 6.90 | 15.72 | 22.13 | 0.36 | | 0.33 | Fibrin | 400 | 620 | 100 | 120 | 1 | | 1 | 55,692 |
| 54.48 | 7.01 | 15.70 | 22.00 | 0.38 | | 0.43 | Egg albumin | 400 | 620 | 100 | 120 | 1 | | 1 | 55,692 |
| 54.84 | 7.09 | 15.82 | 21.23 | 0.68 | | 0.33 | Serum albumin | 400 | 620 | 100 | 120 | 2 | | 1 | 55,893 |

According to Osborne in 1902† (111)

| | | | | | | | | | | | | | | | |
|-------|------|-------|-------|-------|-------|------|----------------|-----|------|-----|-----|---|---|---|--------|
| 52.75 | 7.10 | 15.51 | 23.02 | 1.616 | | | Egg albumin | 696 | 1125 | 175 | 220 | 8 | | | 15,703 |
| 55.23 | 7.26 | 16.13 | 20.78 | 0.600 | | | Zein | 736 | 1161 | 184 | 208 | 3 | | | 15,993 |
| 52.72 | 6.86 | 17.66 | 21.73 | 1.027 | | | Gliadin | 685 | 1068 | 196 | 211 | 5 | | | 15,568 |
| 52.99 | 7.01 | 15.93 | 22.14 | 1.930 | | | Serum albumin‡ | 682 | 1051 | 171 | 207 | 9 | | | 14,989 |
| 54.64 | 7.09 | 17.38 | 20.16 | 0.39 | 0.335 | | Hemoglobin‡ | 758 | 1181 | 207 | 210 | 2 | 1 | | 16,655 |
| 54.57 | 7.11 | 16.38 | 21.03 | 0.568 | 0.336 | | Hemoglobin§ | 758 | 1185 | 195 | 219 | 3 | 1 | | 16,667 |
| 52.68 | 6.83 | 16.91 | 22.48 | 1.10 | | | Fibrin | 645 | 1004 | 178 | 207 | 5 | | | 14,708 |
| 51.50 | 7.02 | 18.69 | 21.91 | 0.88 | | | Edestin | 624 | 1021 | 193 | 199 | 4 | | | 14,523 |
| 53.13 | 7.06 | 15.78 | 22.37 | 0.80 | | 0.86 | Casein | 708 | 1130 | 180 | 224 | 4 | | 4 | 15,982 |

* For a critical study see Laskowsky (Ann. 58, 129 (1846)).

† Osborne quotes the earlier literature in this paper. See, also, Osborne: The Vegetable Proteins, 2nd edition, Longmans, Green and Co. (1924).

‡ Of the horse.

§ Of the dog.

nized to be constituents of protein molecules, as a result of investigations of Kossel (76), Ehrlich (40), Emil Fischer (51), Dakin (36), and most recently of Mueller (98), Rose (120), and Van Slyke (147). Not all proteins contain all known amino acids, and the proportion in which these substances occur varies widely in different proteins. Certain of the amino acids are indispensable for nutrition; certain of them play an important rôle in immunity. Special significance attaches to the sulfur-

containing acids, methionine and cysteine; to the hydroxyl-containing amino acids, serine, threonine, hydroxyproline, and hydroxyglutamic acid, which can form ester linkages with acids; to the phenolic hydroxyl groups of tyrosine; to the hexone bases, histidine, arginine, and lysine, which as free amino acids combine two acid equivalents; and to the dicarboxylic amino acids, aspartic, glutamic, and hydroxyglutamic, which can combine with two equivalents of base per mole when not present as amides.

The characteristic groups of the amino acids must be considered largely responsible for the behavior of the proteins, but the arrangement of these groups, and therefore the structural chemistry of the proteins, cannot yet be considered completely apprehended: The work of Hofmeister and of Emil Fischer leaves no doubt that the amino acids are combined with each other through the peptide linkage. Not all amino acids when bound in peptide linkage are therefore possessed of free amino groups, and not all are possessed of free carboxyl groups, but only those which are trivalent, that is, which have more than one basic, or more than one carboxylic, group. Only the α -amino and carboxyl groups of amino acids are bound in peptide linkage; the side-chain groups—the imidazole, amide, hydroxyl, sulfhydryl, and phenolic groups—are free and give rise to the reactive properties of proteins. The backbone of the protein is presumably that of the peptide chain, with side chains constituting either (1) non-polar hydrocarbon chains, as in alanine, valine, or leucine, (2) imidazole groups, as in histidine, (3) indole groups, as in tryptophane, (4) benzene rings, as in phenylalanine and tyrosine, or (5) strongly polar guanidine, amino, hydroxyl, or carboxyl groups. Regardless of the nature of the side chains, the distance separating them in the stretched peptide is always the same, since it depends only upon the distance of separation between carbon and carbon or carbon and nitrogen atoms. The 3.5 Å. distance separating side chains is the salient repeating line in x-ray analyses of stretched proteins as of peptides. In the case of peptides in solution, however, it has been demonstrated that there is free rotation around carbon-carbon and carbon-nitrogen bonds. The distances separating the reactive groups may therefore be smaller than 3.5 Å. and not much greater than those calculable from a statistical treatment of probable positions on the basis of free rotation (31, 44, 79, 163).

The elementary composition and the densities of proteins are essentially those that could be calculated from the amino acid residues in the molecule. The numbers of free groups and paraffin side chains of proteins could also be calculated, given the amino acid compositions. The amino acid compositions of many proteins may now be considered known, at least as a first approximation. The spatial relations of the free groups to each other could be revealed by organic chemistry, provided more were known regard-

ing the configuration of the peptide chain in the intact protein. Although various theories have been advanced (159), certain knowledge is still lacking regarding the internal structure of proteins. Meanwhile many physical-chemical methods have become available, in terms of which much can be ascertained regarding the size and shape of the molecule and the number and nature of their reactive groups as well as something regarding their spatial distribution.

II. SIZE AND SHAPE OF THE PROTEIN MOLECULE

The large size of the protein molecule was suspected long before accurate methods were available for the determination of molecular weights. Many membranes, both natural and artificial, were known to be impermeable to given proteins, and their sizes could be approximated from the dimensions of the pores of the membranes. Proteins diffuse slowly, and their molecular weights have been estimated from the rate of diffusion. In fields of force many times that of gravity, they can be centrifuged and their molecular weights calculated from the rate of sedimentation. Although the protein molecules are too large—and therefore present in too small concentration—to have significant effects upon vapor pressure or freezing point, their molecular weights can be estimated from measurements of osmotic pressure, under conditions such that membranes are impermeable to the protein but not to the small concentration of impurities from the last traces of which they are freed with difficulty. Oriented by virtue of asymmetry of shape, or of electrical charge, their times of relaxation to random distribution can be determined; and the relaxation time, like the rate of diffusion and of sedimentation, is a function of the size and shape of the molecule.

Analytical methods and the minimal molecular weight

Any accurate determination of a constituent of a protein—provided the protein preparation is pure and of uniform composition—may be employed in calculating the smallest weight that can be ascribed to this protein if it contains 1 gram-atom or gram-molecule of this component. Thus the early determinations of the iron content of hemoglobin led to the calculation that the molecular weight of this respiratory protein could not be smaller than 16,670 if it contained 1 gram-atom of iron. The iron is present in the prosthetic porphyrin group, heme, and it followed therefore that if hemoglobin contained 2 gram-atoms of iron, its molecular weight would be 33,340; if 4 gram-atoms of iron, 66,680. The minimal molecular weight does not, of course, reveal whether all the molecules in solution are of the same size, or what the average molecular weight is if the protein is not monodisperse, even though the

different molecular species have identical composition. If, however, the protein is of uniform composition, the true molecular weight must be a multiple of the minimal molecular weight calculated from any component.

Only a few proteins contain iron, but most proteins contain sulfur. Moreover the sulfur present is of two kinds, as was clearly understood by the analytical chemists of the 19th century such as Schultz (122), Krüger (77), and T. B. Osborne (111). Krüger was the first to present accurate analytical data for the total and the alkali-labile sulfur. Moreover, he pointed out that for egg albumin and for fibrin the ratio of total sulfur to labile sulfur was, respectively, 4:1 and 3:1⁴ (77). The sulfide or labile sulfur therefore represents an integral fraction of the total sulfur. Osborne (111) in 1902 estimated, on the basis of the results in table 1, that the minimal molecular weights of many proteins were in the neighborhood of 15,000 to 30,000,—estimates borne out by recent physicochemical studies.

At the beginning of this century there were no satisfactory analytical methods for determining cystine, and the other sulfur-containing amino acid thus far isolated from a protein hydrolysate, methionine, had not been discovered (98). That part of the total sulfur which is not alkali-labile would appear to represent methionine on the basis of the best analytical measurements at present available (see table 2).

The analyses of the earlier investigators for both total sulfur and alkali-labile sulfur have recently been confirmed for many proteins (10, 17) (table 2). The determination of alkali-labile sulfur is in some cases equal to, in others slightly in excess of, the amount of cystine, or cysteine, estimated to be present. The discrepancy, in the case of the analysis of certain proteins, between alkali-labile sulfur and the sulfur present as cystine or cysteine may represent another thioamino acid, but it may also represent destruction of cystine during hydrolysis (11, 73, 72). Certainly with few exceptions alkali-labile sulfur is now accounted for as cystine (table 2).

Analyses for those amino acids generally present in proteins in small amounts have also been used in the estimation of the minimal molecular weights of the proteins (14, 22, 33). Tryptophane, tyrosine, methionine, and cystine, and occasionally also histidine and arginine, have been particularly useful in this connection, since the analytical methods for their determination are most readily susceptible of being placed on a firm quantitative basis. Improvement in the specificity and the accuracy of analytical methods and in the purity of protein preparations may ultimately be expected to lead to minimal molecular weights and therefore to

⁴ That these are characteristic ratios for the occurrence of amino acids in proteins has since been demonstrated repeatedly (14, 15).

TABLE 2
Composition and molecular weights of proteins

| COMPOSITION IN PER CENT. | | NUMBER OF CONSTITUENTS | | ESTIMATED MOLECULAR WEIGHT | |
|--------------------------|----------------|------------------------|----------------------|---------------------------------|-------------------------------------|
| PROTEIN | | Tyrosine | From common solution | From concentrated urea solution | From sedimentation ultracentrifuge. |
| Total sulfur | Bulidic sulfur | Cysteine | Bulidic sulfur | Tryptophane | Tyrosine |
| 0.49 | 1.22 | 5.24 | 1.18 | 4.00 | Egg albumin |
| 1.60 | 1.33 | 5.07 | 1.28 | 4.21 | |
| 3.34 | 12.5 | | 12.2 | Insulin* | |
| † | 0.21 | 0.91 | 2.58 | 5.66 | Zein |
| | 0.22 | 2.40 | 2.25 | 5.90 | |
| 1.03 | 0.62 | 2.15 | 1.60 | 0.83 | Gliadin |
| 0.99 | 2.40 | 2.03 | 1.82 | 3.50 | |
| 0.39 | 0.19 | 0.41 | 1.28 | 3.15 | Hemoglobin (horse) |
| | 0.18 | | | | |
| † | † | 6.02 | 0.52 | 4.66 | Serum albumin |
| | | 5.80 | 0.26 | 4.79 | |
| 0.76 | 0.10 | 0.21 | 3.25 | 1.54 | Casein |
| 0.80 | 0.09 | 0.26 | 3.36 | 1.60 | |

| | | | | | | | | | | | | |
|------|------|------|------|------|--------------------|---------|----|---|----|---------|-----------|---------|
| 1.10 | 0.62 | 3.4 | 0.82 | 3.38 | Myosin (rabbit) | 3 | 24 | 4 | 18 | 100,000 | 1,000,000 | |
| | 0.77 | | | | | | | | | | | |
| 0.88 | 0.35 | 1.22 | 2.30 | 1.45 | 4.53 | Edestin | 6 | 6 | 8 | 8 | 13 | 54,000 |
| 0.90 | 1.35 | 2.40 | 1.56 | 4.58 | | | 9 | 9 | 9 | 14 | | 309,000 |

* Crystalline insulin contains 0.52 per cent zinc (123). Assuming three zinc atoms per insulin molecule leads to a molecular weight of 37,700.

† The values reported are so discrepant that they are not considered here.

‡ Based on the iron analysis.

§ In ethanol-water mixtures.

Note added in proof: Kuhn, Birkofe, and Quackenbush (Ber. 72, 407 (1939)) have published a new study of sulfur distribution in proteins since this table was compiled. For egg albumin their value for the total sulfur, 1.59 per cent, is in good agreement with Osbome's, and this and their somewhat lower value for methionine, 4.85 per cent, leads to a molecular weight of 36,000 to 37,000 on the basis of twelve residues per mole. The yield of cystine, 1.78 per cent, is moreover far greater than that previously reported, yielding six cysteine residues and removing the discrepancy noted above between sulfide sulfur and cysteine.

Their values for insulin suggest one methionine residue per mole and are therefore in fair agreement with the results of Kassel and Brand (72). It should be pointed out however that Kuhn's study was on amorphous and not on crystalline insulin. They also studied horse hemoglobin and globin and report values of methionine and cystine suggesting two residues of the former and three of cysteine on the basis of a molecular weight of approximately 33,000. Five atoms of total sulfur are thus demanded by their analysis, as compared with the four previously assumed, and three atoms of sulfide sulfur as compared with the two previously assumed.

The only other protein in the above list studied by Kuhn, Birkofe, and Quackenbush is casein. As in the case of egg albumin and hemoglobin their cystine value, 0.42 per cent, is higher than those tabulated above, removing the discrepancy between cysteine and sulfide sulfur, and their methionine is smaller, namely, 3.03 per cent.

Sulfur and sulfide sulfur.—The earlier literature has been considered by Osborne (111) and in our first study of minimal molecular weights (33). The more recent studies on labile sulfur, the results of which are in many cases in excellent agreement with Osborne's values, are by Clarke and his associates (17, 164), Baernstein (10), and Bailey (11). The relation between total sulfur and cystine sulfur in *insulin* has been considered by Miller and du Vigneaud (95). Assuming thirty-six atoms of sulfur per mole of insulin leads to a molecular weight of 34,488. The cystine analysis suggests that most if not all of the sulfur of insulin is sulfide sulfur.

The new analyses of Blumenthal and Clarke, giving 0.22 per cent of alkali-labile sulfur in *zein*, are in excellent agreement with Osborne's data. Baernstein's value of 0.99 per cent for the total sulfur in *gliadin* is in good agreement with Osborne's early value, but Bailey (11) reports a higher result, 1.19 per cent. Blumenthal and Clarke's result of 0.09 per cent for the alkali-labile sulfur of *casein* is far lower than Osborne's value of 0.76, which is in better agreement with Baernstein's value of 0.80 per cent (10) and Kassel and Brand's value of 0.78 per cent (73). The analyses for sulfur of rabbit *myosin* are due to Bailey (11, 12). Osborne's early value for the sulfur in *edestin*, 0.88 per cent, is slightly lower than the values 0.93 reported by Bailey (11), 0.97 reported by Kassel and Brand (73), and 0.99 reported by Baernstein (10).

Cystine.—The earlier literature, including the work of Folin and Sullivan and their associates, is summarized in table 19 of reference 28. Since then Sullivan has modified his procedure, and Vickery and White (150) have considered the cystine yielded by proteins, as have Baernstein (8) and Bailey (11). The higher value given for *egg albumin* is from Calvery (22). The cystine content in *insulin* is from Miller and du Vigneaud (95). The value 0.91 per cent for *zein* is from Vickery and White (150) and is intermediate between the earlier values quoted (28). Vickery and White's average value for *gliadin* is lower than most values in the literature, but 2.5 per cent is in good agreement with the previous values, although the value is high and is higher than the value 2.40 per cent, which has also been reported (11, 28, 150). The cystine determination on horse *hemoglobin* is by Vickery and White (150). Hewitt reports a cystine content of 5.80 per cent for *serum albumin* (64), a value intermediate between Sullivan's early value of 5.71 per cent (134) and Folin and Marenzi's value of 6.02 per cent (53). The early literature on the cystine content of *casein* is consistent with a large number of values falling between 0.22 and 0.30 per cent (28). Vickery and White's new determinations fall within the same limits, 0.20 and 0.24 per cent. There thus appears to be a discrepancy between the sulfide sulfur and the cystine analyses. This determination renders difficult the estimation of the molecular weight in the neighborhood of 33,000 as demanded by osmotic pressure in urea. Vickery and White (150) have made a large number of analyses on the cystine in *edestin* and their results, which are not far from those of Sullivan and of Folin and Marenzi (28), give an average of 1.25 per cent. Baernstein's values range from 1.14 to 1.36 per cent.

Methionine.—The methods for determining the per cent of methionine in proteins and the values obtained have been considered by Baernstein (9, 10,), Blumenthal and Clarke (17), Bailey (11), and Kassel and Brand (73), and have been reviewed by Toennis (144). In the case of *zein* Baernstein (9) reports values varying from 2.25 to 2.58 per cent, depending upon the source of the preparation and the method of analysis. His value for *gliadin* is far higher than Bailey's. Baernstein (9) reports a value of 3.53 per cent for a *casein* preparation the origin of which he does not give. One supplied by R. A. Gortner had a methionine content of 3.36 per cent and one by

D. B. Jones a methionine content of 3.25 per cent, whereas a Harris preparation, studied by the volatile iodide method by Baernstein (9), had a value of 3.21 per cent and by Kassel and Brand (73) a value of 3.2 per cent. The methionine values for *edestin*, as reported by Baernstein, Bailey, and Kassel and Brand, all run between 2.3 and 2.4 per cent.

Tryptophane and tyrosine.—The earlier literature, including the work of Folin and his collaborators and of Linderström-Lang on casein, has previously been summarized (table 19 of reference 28). Not included in this summary are studies on *egg albumin* by Calverly (22), whose estimate for tryptophane is 1.28 per cent and for tyrosine 4.21 per cent, and on *insulin* by du Vigneaud (151).

On the basis of the tryptophane estimates of a number of investigators the molecular weight of *zein* would have to be many times that observed both by osmotic pressure and ultracentrifugal measurements. No more than 0.2 per cent tryptophane has ever been reported for *zein*, and a large number of investigators report that *zein* contains no tryptophane.

The fractions of *gliadin* separated by Haugaard and Johnson (62), although constant in their histidine content, which was close to that reported by Van Slyke (145), vary appreciably in the percentage of certain other of the residues studied. Their least soluble fraction, IV, had a tryptophane content of 1.08 per cent, in excellent agreement with a number of previous determinations (28, page 861). The tryptophane contents of their other fractions were greater than of their fraction IV, and greater the greater the solubility and the greater the acid-insoluble humin nitrogen, reaching a value of 1.82 per cent, or four residues in a 41,000 molecular weight fraction. Conversely, the tyrosine content of their least soluble fraction was highest, 3.02 per cent, and close to that given by Looney (87), 3.04 per cent.

The tryptophane content of *casein* has repeatedly been studied; most results fall between 1.4 and 1.6 per cent, the higher value being due to Linderström-Lang (84). The values for tyrosine reported have been too discrepant to consider. The tryptophane results reported for *edestin* generally range from 1.45 to 1.56 per cent and for tyrosine from 4.53 to 4.58 per cent (28).

The tryptophane content reported by early workers for *serum albumin* was always high, perhaps because of the difficulties in purification, since serum albumin can be separated into a number of crystalline fractions of different solubility (131), different dipole moment (49), different carbohydrate content (84, 131), and different tryptophane content. Fürth and Lieben (56) reported 1.3 per cent, and Hunter and Borsook (69) 1.79 per cent of tryptophane recovered from serum albumin. Folin and Marenzi (68) studied a preparation of serum albumin that had been recrystallized from five to seven times in this laboratory and found 0.52 and 0.53 per cent of tryptophane in the fractions studied. More recently Hewitt (84) has studied a serum albumin recrystallized twelve times and obtained a value just half of that reported by Folin and Marenzi (58), namely, 0.28 per cent.

Although tryptophane appears to be a variable in the various fractions of serum albumin, the tyrosine content would appear to be invariant. Hunter and Borsook (69) reported 4.63 per cent, Folin and Marenzi (58) 4.66 and 4.67 per cent, and Hewitt's crystalalbumin preparation yielded 4.79 per cent. Assuming a molecular weight of 73,000 and eighteen tyrosine residues leads to a value of 4.47 per cent tyrosine per mole, which is rather lower than that observed. On the basis of eighteen tyrosine residues, these results thus suggest a molecular weight of between 68,000 and 70,000 rather than 73,000.

Osmotic pressure.—Sörensen estimated the molecular weight of *egg albumin* as 34,000 (130); this value was confirmed by Burk and Greenberg (20) in aqueous solution and by them and Huang and Wu (66) in isoelectric urea solutions. Marrack and Hewitt (89) estimated the molecular weight to be 43,000 at 37°C., but Nichols (105) attributes this higher value to the presence of aggregating material.

Burk (19) has estimated the molecular weight of *zein* as 39,000 in alcohol-water mixtures and 37,000 in urea. I am indebted to Dr. Burk for recalculating the former value, correcting for deviations from the ideal solution law and for the Donnan effect; the value so obtained was 39,000. Burk has also estimated the molecular weight of *gliadin* as being 41,000 in alcohol-water mixtures and 44,200 in urea (19). Adair (1, 2) estimates the molecular weight of *hemoglobin* to be 67,000 in aqueous solution, but Burk and Greenberg (20) and Wu and Yang (160) (for hemoglobin of the horse but not of the cow) found a value approximately half of this, namely, 34,300, in urea solution. Adair and Robinson's (4) estimate for the molecular weight of serum albumin is 72,000, that of Burk (18) 74,600, that of Roche and Marquet (119) 69,000; that of Sörensen, as corrected by Burk (18) and by Adair and Robinson (4) for measurements in 5 per cent ammonium sulfate solution, is 76,300. Burk and Greenberg (20) give the molecular weight of *casein* in urea solution as 33,600 and of *edestin* in this solvent as 49,000 (20). The molecular weight of *myosin* is of the order of one million (153). According to Weber, who has also studied the molecular weight of myosin in urea solution (154) with the method of osmotic pressure, myosin is not so large in this solvent and has a molecular weight of close to 100,000.

Ultracentrifuge.—The new values for *egg albumin* from Upsala are 43,800 with the sedimentation and 40,500 with the equilibrium ultracentrifuge. We are indebted to Professor Svedberg for informing us of the unpublished results of Pedersen (116; also (82). Williams and Watson (157) have studied the molecular weight of egg albumin in urea solution and report a molecular weight in this solvent of approximately 21,000. The sedimentation value for *insulin* is 40,900 and the equilibrium centrifuge value 35,100 (127). The sedimentation value for *zein* is 35,000 (152), for *gliadin* 26,000 (6), and for the *carboxyhemoglobin* of the horse 69,000. The estimate by means of the equilibrium ultracentrifuge for the latter protein is 68,000. The sedimentation value for horse *serum albumin* is 70,200 and the value from the equilibrium centrifuge is 66,900 (102, 116). Svedberg, Carpenter, and Carpenter (136) find that within the limits of experimental error the acid-alcohol soluble portion of *casein* was homogeneous with regard to molecular weight and that therefore it probably was a pure chemical individual. The molecular weight was found to be $375,000 \pm 11,000$ (137), though Hammarsten casein as a whole was found to consist of protein molecules of different molecular weight. Svedberg, Carpenter, and Carpenter later examined casein prepared by the method of Van Slyke and Baker (146) and found that the bulk of the crude casein prepared by this method had a molecular weight between 75,000 and 100,000 (137, page 710). The new value from sedimentation and diffusion for the molecular weight of *edestin* is 309,000, or approximately six times the weight reported from the osmotic pressure in urea.

true molecular weights of very great accuracy.⁵ This is especially true of the beautiful new methods being developed by Bergmann (14).

⁵ In the case of smaller molecules physical-chemical methods, such as boiling points and freezing points, yield the number of times the empirical formula weight must be multiplied to give the true molecular weight.

Analytical methods which enable a distinction to be made between —SS— and —SH groups in the intact protein have also been developed in recent years (59, 65, 80, 96), and it would appear that reagents such as urea and guanidine, which bring about changes in molecular weights of certain proteins, in many cases also influence these configurations. Here, too, the simultaneous use of analytical and physicochemical methods may well lead to significant advances in our understanding of the structure of the protein molecule.

Osmotic pressure methods and the average molecular weight

Measurements of osmotic pressure yield the average molecular weight, provided adequate correction is made (1) for the unequal distribution of ions across the membrane due to the Donnan equilibrium and (2) for the osmotic coefficients of the protein in the solution. In practice the Donnan effect is largely eliminated by carrying out the osmotic pressure measurements not far from the isoelectric point in salt solutions, generally in buffer solutions. But osmotic coefficients deviate appreciably from unity at finite protein concentration, and molecular weights estimated from osmotic pressure measurements therefore involve extrapolation from a series of measurements in which protein concentration is varied. If pH is also varied, so that the protein has a net charge, the osmotic pressure increases with the valence of the protein ion, and estimation of the molecular weight under these circumstances is still more complicated.

Although Hufner and Gansser (67) attempted to determine the molecular weight of hemoglobin by means of osmotic pressure measurements, it remained for Sörensen (130) adequately to consider the correction necessary for the interpretations of such measurements if they were to yield molecular weights. His study of egg albumin, published just twenty-one years ago, marked a turning point in protein chemistry and suggested that the molecular weight of this protein was in the neighborhood of 34,000. Sörensen has since studied other proteins, and at least two other investigators, Adair (1, 2) and Burk (18, 19, 20), have made large numbers of adequate studies of molecular weights of proteins under a variety of conditions by the osmotic pressure method. The disadvantage of the osmotic pressure method is that at best it can yield only an average molecular weight; its advantage is that it is a thermodynamic measurement, dependent only on the number of particles in solution and not on their shape.⁶

⁶ As we shall subsequently see, the activity coefficient, and therefore the osmotic coefficient due to the interaction of protein and electrolyte, are presumably functions of the valence of the latter and of the size and shape and electric moments of the former. The form of the extrapolation for the determination of osmotic pressure may thus also reflect the shape of the protein molecule.

In the course of osmotic pressure measurements Burk and Greenberg (20) noted in 1929 that the molecular weight of horse hemoglobin in concentrated urea solution, instead of being in the neighborhood of 66,680, as in aqueous solution, was half this value, or in the neighborhood of 33,340. This observation, confirmed by Wu (160) for the hemoglobin of the horse but not of the cow, and by Steinhardt (133) in studies at Upsala by means of the ultracentrifuge, has since been extended to a variety of other proteins, many of which are broken down into molecules of smaller size in the presence of this solvent (18, 19, 66, 154, 157). A series of other solvents, including other amides, but also guanidine, certain amino acids, and proteins (91, 113, 115), also possess this property and open a new chapter in protein chemistry. Whether or not these changes in molecular weight are to be considered as "dissociations" or as destruction of "aggregation," and whether or not they are reversible and therefore reveal the state of the native protein, they suggest that certain of the forces binding vast protein molecules together are different from others, and thus they add greatly to our knowledge regarding the structure of the proteins.

Sedimentation methods and the molecular weight

If the molecular weights of proteins were as large as analytical and osmotic pressure measurements suggested, it followed that they would be sedimented from solution by sufficiently great centrifugal forces. It remained for Svedberg (140) in 1926 to develop in the ultracentrifuge a tool adequate to the task. This most important tool, the various improvements it has undergone, and many of the results that have been obtained with it, as well as the equations employed in the calculation of molecular weights from ultracentrifugal measurements, have been reported in two previous contributions to this Journal by Svedberg (135).

The great strides made possible by this development depend in no small part upon the optical systems devised, which render it possible to distinguish more than one boundary and therefore more than one sedimentation rate, and thus to determine the molecular weight of more than one component in solutions that are not monodisperse, that is, that consist of proteins of more than one size. Thus far the method has been used largely to prove that the protein preparations under investigation consisted of molecules uniform with respect to size and to determine the molecular weights of the various proteins in certain natural fluids, such as milk (114) and the blood of certain invertebrates (42, 139). The fact that not only solvents, such as urea and arginine, but also certain proteins may influence each other's sedimentation rate (91) opens a new vista for the determination of the inter- and intra-molecular forces involved and also the possibility of apprehending more regarding the state of proteins in nature.

One cannot overemphasize the great advances that have thus far been achieved and that can be expected in the future from ultracentrifugal studies. The interpretation of the experimentally observed sedimentations involves a theory to which contributions are still being made (82, 116). The equation relating the molecular weight to the sedimentation constant includes a term for the frictional resistance experienced by the molecule. This term, which also appears in the Sutherland-Einstein derivation of the translational diffusion coefficient, is a function of both the size and the shape of the molecule. The combination of the sedimentation constant and diffusion coefficient permits the elimination of the shape factor from molecular weight calculations.⁷ Sedimentation velocity measurements alone would give different molecular weights for molecules of the same composition and weight but different shapes. Conversely, molecules of totally different composition, but of the same size, density, and shape, would appear to be monodisperse in the ultracentrifuge if their isoelectric points were not too different. The latter statement has come to have more than theoretical significance, because so many proteins have molecular weights ranging from 34,000 to 42,000 and isoelectric points ranging from 4.7 to 5.7. Moreover, molecules of different sizes and shapes might conceivably have the same sedimentation constant. To deduce identity in the chemical nature of a protein from identity in its sedimentation velocity is a danger clearly envisaged by all those familiar with the methods.

Diffusion methods and the size and shape of the molecule

The slow diffusion of the proteins was one of the early observations which led to their classification as colloids. Determined with less elaborate apparatus than that demanded for ultracentrifugal measurements, the diffusion coefficient has also been used in estimating molecular weights (82, 106).

In one simple form of diffusion apparatus the concentration gradient is within the interstices of a sintered-glass filter, the pores of which must be large in comparison with the diameter of the protein molecule, but small enough to prevent mixing by convection (106). The diffusion coefficient

⁷ Polson (118), using a combination of the viscosity equations of Kuhn and Arrhenius, has attempted, by inserting the asymmetries deduced from viscosity measurements in the equation of Herzog, Illig, and Kudar (63) for the diffusion ratio between a spherical and an elongated ellipsoidal particle, to determine the molecular weights of proteins. The results are of the right order of magnitude, and differ by a constant from the best results obtained by the sedimentation method. It would seem, however, that these were known with far greater certainty than the proper form of the viscosity equation. The influence of shape is also eliminated in measurements with the equilibrium centrifuge.

may be calculated from the changes with time of the concentrations on the two sides of the disc, provided the latter has been calibrated with a substance of known diffusion coefficient. If the analysis is specific for a given protein, this method may be employed in the presence of other proteins; if non-specific, measurements upon mixtures of proteins are difficult of interpretation, since they yield diffusion coefficients which are not averages but are weighted in favor of the smaller and more rapidly diffusing protein components.

The optical methods perfected at Upsala for the study of concentration gradients have been applied to diffusion measurements (82) and have greatly improved the classical method of free diffusion. The disturbances occasioned by sampling are avoided, and a more complete analysis of the diffusion is possible. As in the ultracentrifuge, the data reveal whether or not more than one component is present. The study of mixtures presents many difficulties, however, unless the optical method be made specific for a particular molecular species. This method, in contrast to that involving sintered-glass filters, yields not relative but absolute values of the diffusion coefficient.

The diffusion coefficient is a function of the size of the molecule. For the case of spherical molecules the diffusion coefficient is, according to theory, inversely proportional to the radius. A small error in the diffusion coefficient is thus greatly magnified in the estimation of the molecular weight.

The diffusion coefficient is also a function of the shape of the molecule. An ellipsoid encounters more frictional resistance than a sphere of the same volume, and its diffusion coefficient is accordingly smaller. All proteins have in fact smaller diffusion coefficients, D , than those calculated, D_0 , from their molecular volumes on the basis of spherical shape. The ratio D_0/D , equal to the Svedberg dissymmetry constant (135), may be interpreted in terms of either departure from spherical shape, or hydration, or both. If the hydration be assumed zero, the equations of Herzog, Illig, and Kudar (63) or of Perrin (117) may be used to calculate from D_0/D the ratio of major to minor axes (s/d) of an ellipsoidal model for the molecule. Such calculations have been reported by Polson (118) and by Neurath (104).

Although the Svedberg dissymmetry constant is often only a little greater than unity, when the values of D and D_0 are substituted in the Perrin equation the asymmetries suggested are very large. Thus, according to this treatment, the dissymmetry constant of 1.18 for egg albumin yields a ratio of major to minor axis of 4:1; thyroglobulin with a dissymmetry constant of 1.41 has, on the same basis, a ratio of major to minor axis of 8:1.

Dielectric constant dispersion methods and the size and shape of the molecule

Proteins, like other molecules, may be expected to assume a random distribution in the body of a dilute solution⁸ as a result of thermal motion. All proteins that have thus far been investigated are, however, oriented in an electric field. By virtue of their structure as dipolar ions they follow an alternating potential, provided its frequency is not too great. At low frequencies proteins thus increase the dielectric constants of solutions. At high frequencies, where they are unable to follow the alternations, the

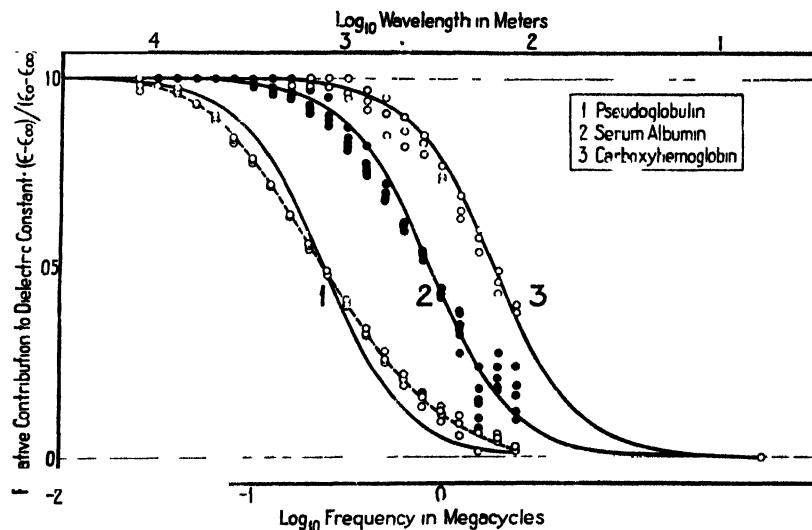


FIG. 1. Dispersion curves of blood proteins (horse)

dielectric constant of a solution is decreased, for the vast protein molecules displace solvent dipoles which could have oriented in the electric field (107, 109).

Plotting the dielectric constant of a solution containing protein against the frequency of the alternating potential yields a dielectric constant dispersion curve, which is characteristic of the protein. The theory of such curves has been developed by Debye (39). The methods that have been employed and the dielectric constants that have been observed have recently been reviewed elsewhere (31, 41, 109, 157). The first dispersion

* At the surface of a solution the forces must be considered different, and orientation and layering of the protein result, generally with compression and distortion of the native molecule. The problems associated with surface phenomena are discussed in the previous paper in this Symposium and are therefore not considered here (34). In sufficiently concentrated solutions, or in solutions of low dielectric constant, orienting forces may also be expected to interfere with a random distribution.

curve for a protein was that of Wyman on zein (161). The relation of such curves to the molecular weight has been considered by Wyman, Williams, Arrhenius, Oncley, and Ferry (6, 7, 49, 107, 156, 157, 161, 162).

The extent to which the dielectric constant of a solution is raised by a protein solute depends upon the electrical asymmetry of the molecules; it is a function of their dipole moment, which is determined by the distribution of the electrically charged groups on their surfaces, and has been considered in connection with the number and distribution of the electrically charged groups of proteins (31). The shape of the dielectric constant dispersion curve and the frequencies at which it occurs are, however, functions of the size and shape of the molecule and are therefore considered here.

Studies upon three proteins are graphically represented in figure 1, the ordinate of which is so defined that when the protein molecules are unable to follow the alternating field and contribute nothing to the dielectric constant, its value is zero; when they rotate in phase with the field and make a maximum contribution to the dielectric constant, its value is unity. The curve for hemoglobin of the horse (107) closely follows the form given by Debye's theory. From the critical frequency, ν_c , the relaxation time ($= 1/2\pi \nu_c$) is found to be 8.4×10^{-8} sec. For comparison, we may calculate the time of relaxation expected for a spherical molecule of molecular weight 66,670 and specific volume $v = 0.75$ rotating in a medium with the viscosity (η) of water as $\tau = 3\eta Mv/RT = 5.4 \times 10^{-8}$ sec. The discrepancy may again be interpreted in terms of either hydration or spatial asymmetry. If we assume no asymmetry and postulate an amount of hydration equal to that estimated by Adair and Adair (3) for hemoglobin crystals, and assume that the water of hydration rotates with the protein molecule, the calculated relaxation time is 7.6×10^{-8} sec., in fair agreement with that observed. The same amount of hydration would closely account for the observed diffusion, as an alternative interpretation to an elongated ellipsoidal shape with an axial ratio of 4. The latter shape is not consistent with the single relaxation time observed, unless the dipole moment be perpendicular to the long axis. On the other hand, calculations by Mehl and Ferry show that both D_0/D and the observed relaxation time would be given by a postulate of no hydration and the shape of a flattened ellipsoid of revolution with an axial ratio of 4. For such a body the relaxation times about the different axes are nearly equal, and even if the dipole were inclined to all axes, the dispersion would be indistinguishable from a simple Debye curve.*

* The hemoglobin of the pig has the same molecular weight as that of the horse, but the dielectric constant measurements of Arrhenius (7) yield a relaxation time about twice as great, and therefore exceeding much more (by a factor of three) the value calculated from the molecular weight.

The molecular weight of the crystalline serum albumin of the horse is very close to that of horse hemoglobin, being estimated by Svedberg as 67,000 on the basis of ultracentrifugal measurements (135), and by Adair (4) as 72,000 and by Burk (18) as 74,600 on the basis of osmotic pressure measurements. The difference between the relaxation times of these two proteins would therefore be scarcely perceptible on the basis of their molecular weights, and no sufficiently great differences in hydration to account for the observed dispersion curves can be reasonably postulated. The curves of both proteins appear to follow Debye's theory, but the relaxation time obtained for serum albumin, 19×10^{-8} sec., is over twice as great as that for hemoglobin or more than three times that calculated from its molecular weight. Moreover, the viscosities of serum albumin solutions are far greater than those of hemoglobin solutions at the same volume fraction (37, 45, 82). This demonstrates how relaxation times reflect not only the size but also the shape of proteins. Dielectric constant dispersion curves may therefore become of even more use than sedimentation constants in characterizing many proteins, since so many proteins appear to have the same molecular weights. The relaxation time of a protein may readily be employed in calculating its rotary diffusion constant.

Horse serum pseudoglobulin has a far greater effect upon the dielectric constant, reflecting greater electrical asymmetry than either the serum albumin or the hemoglobin of the horse. As judged by viscosity measurements (25, 26, 45), it is also a less spherical molecule. Its molecular weight in aqueous solution has been estimated to be 150,000 or 162,000 (91, 103). The dielectric constant dispersion curve of this protein differs in two respects from the others in figure 1. In the first place, it occurs at much smaller frequencies; in the second place, its shape does not conform to Debye's theory. Analyzable in terms of two or more relaxation times, this curve suggests either that more than one molecular species was present or that the asymmetry of the molecule is sufficient to reveal relaxation times characteristic of rotation about its different axes. Regardless of the interpretation of such a curve, however, it further illustrates the usefulness of dielectric constant studies for the characterization of proteins and for demonstrating differences and similarities between similar proteins from different sources.¹⁰

The relaxation times of those proteins which appear in figure 1 are summarized in table 3, together with some others. There are as yet no satisfactory experimental estimates of relaxation times of small molecules, such as water, amino acids, and peptides; the order of magnitude to be expected is, however, given by values calculated from the molecular volumes.

In the case of zein, beautifully studied by Williams and his coworkers

¹⁰ For instance, in the comparison of blood and tissue proteins or blood and urinary proteins (55).

TABLE 3
Relaxation times calculated from molecular volumes, from dispersion of the dielectric constant, and from double refraction of flow

| SUBSTANCE | MOLECULAR WEIGHT <i>M</i> | MOLECULAR VOLUME $\frac{V}{M}$ | RELAXATION TIME | |
|-------------------------------------------|------------------------------|-----------------------------------|------------------------------|-------------------------------|
| | | | $\frac{3\pi V^*}{RT}$ | $\frac{1}{2\pi c_0}$ |
| Water..... | 18 | 18 | 2×10^{-11} | $32 \times 10^{-8} \dagger$ |
| Glycine..... | 75 | 57 | 6×10^{-11} | 53×10^{-8} |
| Diglycine..... | 132 | 93 | 10×10^{-11} | 8.3×10^{-8} |
| Zein (in 72 per cent ethanol) § (41)..... | 39,000 | 29,000 | 6.9×10^{-8} | |
| | | | $58 \times 10^{-8} \dagger$ | |
| | | | $9.2 \times 10^{-8} \dagger$ | |
| Gliadin (in 60 per cent ethanol) (6)..... | 42,000 | 31,000 | 10×10^{-8} | 9×10^{-8} |
| Lactoglobulin (50, 124)..... | 40,000 | 30,000 | 3.3×10^{-8} | 7×10^{-8} |
| Egg albumin (108, 124)..... | 34,000 | 25,500 | 2.8×10^{-8} | $12 \times 10^{-8} \dagger$ |
| | | | $17 \times 10^{-8} \dagger$ | 20×10^{-8} |
| | | | $3.7 \times 10^{-8} \dagger$ | 4.4×10^{-8} |
| Hemoglobin (horse) (43, 107)..... | 66,700 | 50,000 | 5.4×10^{-8} | 8.4×10^{-8} |
| Serum albumin (horse) (49)..... | 73,000 | 54,000 | 6.8×10^{-8} | 19×10^{-8} |
| Serum pseudoglobulin (horse) (49)..... | 162,000 | 120,000 | 13×10^{-8} | $66 \times 10^{-8} \dagger$ |
| | | | $17 \times 10^{-8} \dagger$ | 33×10^{-8} |
| | | | $79 \times 10^{-8} \dagger$ | 132×10^{-8} |
| Tobacco mosaic virus (94)..... | 50,000,000 | 38,000,000 | 4.1×10^{-8} | $(1 \times 10^{-2}) \ddagger$ |
| Myosin (cow) (94, 100)..... | | | | $(5 \times 10^{-3}) \ddagger$ |

* On the assumption of spherical shape, without correcting for hydration.

† On the assumption of elongated ellipsoidal shape, without correcting for hydration, using the axial ratio given by the ratio of the observed relaxation times.

‡ From the single average critical frequency which best describes the dispersion.

§ The earlier measurements of Wyman (161) in 70 per cent propanol give an average relaxation time of 6.2×10^{-8} sec., as against a calculated value of 10×10^{-8} .

¶ Estimated from double refraction of flow.

(41, 157), and of serum pseudoglobulin and egg albumin, studied by Oncley, Ferry, and Shack (49, 50, 108, 124) in this laboratory, the dispersion, although it can be roughly described by a single "average" critical frequency, is much better represented by two different frequencies which yield different relaxation times. If these be interpreted in terms of rotation of an ellipsoid about different axes, then the asymmetry may be evaluated from the ratio of the relaxation times, *without any arbitrary assumptions concerning hydration.*

The asymmetries, which are themselves independent of hydration, may then be used, with the assumption of zero hydration, to calculate the absolute magnitudes of the different relaxation times from the molecular volumes of the respective ellipsoids. These figures are included in table 3, and for zein and egg albumin are in good agreement with the experimental relaxation times, suggesting that for these proteins hydration has little influence on the rotary diffusion.

Double refraction of flow and the size and shape of the molecule

The phenomenon of double refraction of flow of protein solutions has been discussed so recently (94) that only its use in the study of the size and shape of the molecule need be mentioned here. When ordinary light is passed through a Nicol prism, it is plane polarized and can be extinguished by a second Nicol prism the plane of polarization of which is at right angles to that of the first. If a velocity gradient is produced in a solution of asymmetrical molecules placed between the prisms, light will again pass, as the result of double refraction resulting from the orientation of the molecules. This effect will be greater the more asymmetrical the molecules and the greater the velocity gradient to which they are subjected, that is to say, the larger the proportion of molecules oriented in the stream lines of the solution. From measurements of the double refraction the extent and the principal direction of orientation of the molecules may be readily determined.

The orienting forces in the solution may be considered as opposing disorienting rotatory Brownian motion. Calculations of resistance to rotation of ellipsoidal particles in a viscous medium have been made by Gans (57) and Perrin (117), and more roughly by Werner Kuhn (78). It follows from their calculations that the relaxation time of a very long ellipsoidal particle, rotating about its shorter axis, is approximately proportional to the cube of its length. On the basis of their relations and the studies of Edsall and von Murali upon myosin (100) it has been estimated that this muscle globulin has a length in the neighborhood of 8500 Å. (94). This value may be compared with estimates of length calculated on the basis of the molecular volume and estimates of the asymmetry, s/d , for less asymmetric proteins. Tentative calculations of this kind yield lengths for egg

albumin of 120 Å., for horse serum albumin of 160 Å., for horse serum globulin of 290 Å., and for tobacco mosaic virus of 5200 Å.

Collagen (94), fibrinogen, ovoglobulin (21), thyroglobulin, tobacco mosaic, and certain other viruses, and the hemocyanins of *Helix* and *Limulus* (88) have also been shown to exhibit double refraction of flow. At higher velocity gradients it is to be expected that less asymmetrical and smaller molecules will also exhibit this phenomenon, and we are at present engaged in reconstructing our apparatus at Harvard University in order to estimate the size and shape of other proteins by this method.

The relaxation times of the asymmetrical molecules thus far studied, calculated as rotation about an axis perpendicular to the longest axis of the molecule from measurements of double refraction of flow by means of Boeder's treatment (94), are given in table 3. For myosin and tobacco mosaic virus they are respectively 1×10^{-1} to 2×10^{-2} . Assuming a molecular weight for this virus of 50,000,000 and calculating its relaxation time on the assumption of spherical symmetry would lead to a relaxation time of 3×10^{-6} or approximately one thousand times smaller than that observed (94).

In table 3 comparison is also made between the relaxation times calculated from dielectric constant dispersion curves. Those observed thus far range from 10^{-8} to 10^{-6} sec., and in every case in which the solvent was water were greater than relaxation times calculated from molecular weights on the basis of spherical symmetry. Extension of the dielectric constant method to far longer wave lengths and of the double refraction of flow method to far greater velocity gradients should render it possible ultimately to study the same protein by these two quite different methods—one electrical, one optical—and thus acquire more satisfactory knowledge regarding the forces involved in the orientation of protein molecules.

Viscosity and the shape of the molecule

According to Einstein's theory, the viscosity of a suspension or solution of uncharged, incompressible spheres is independent of their size, provided they occupy the same volume fraction of the solution and provided their size is large compared with that of solvent molecules. Experimentally it may be shown that when egg albumin and glycine occupy the same volume fraction, the viscosities of their solutions are somewhat similar (37). The radius of the protein calculated as a sphere is approximately 22 Å. and that of the amino acid is 2.8 Å., but this difference in size has little influence on viscosity. In the amino acid series, however, we have shown that dilute solutions of the more asymmetrical molecules, such as β -alanine and ϵ -aminocaproic acid, are far more viscous, and the more so the greater the number of carbon atoms between the positively charged ammonium and the negatively charged carboxyl group (37).

Proteins differ widely from each other with respect to the viscosity of their solutions. Harriet Chick (25, 26) reported a large number of viscosity studies,¹¹ as did Loeb (85) and many subsequent investigators. In Miss Chick's studies serum albumin solutions were always more viscous than egg albumin, and pseudoglobulin and euglobulin still more viscous. The general magnitude of her results has been confirmed (27, 37, 45), and we have assumed therefore, in the light of our studies upon amino acids and peptides of known structure and of Staudinger's studies on other molecules (132), that this was the order in which the shapes of these molecules deviated from spherical symmetry. Deviation between calculated and observed relaxation times is also greater for serum albumin than for hemoglobin, and for pseudoglobulin than for either of these. This is the order of their viscosities.

Although there would appear to be no doubt from these various experimental studies that viscosity is a most sensitive index of asymmetry, a completely satisfactory theoretical equation relating these two properties remains to be developed. The problem has been repeatedly considered (60, 61, 68, 78, 132), and various extensions of the Einstein equation have been suggested and tentatively employed. Estimates of the relation of the two principal axes of the molecule must be adopted with caution, until they prove to be identical with those derived from diffusion measurements, dielectric constant measurements, and measurements of double refraction of flow.

All of these measurements are technically difficult, and the assumptions involved in their theoretical interpretation include such approximations as that molecules are rigid and incompressible, that there is no interaction between solute molecules, and that the solvent may be treated as a structureless continuum. That the degree of asymmetry for the series of proteins thus far studied by such different methods is so nearly of the same order nevertheless suggests that our knowledge begins to be satisfactory regarding not only the size but also the shape of the protein molecule.

III. NUMBER AND DISTRIBUTION OF THE ELECTRICALLY CHARGED GROUPS OF THE PROTEIN MOLECULE

Analytical methods and the number of reactive groups

Proteins as a class are extremely diverse, not only in their biological functions but also in their chemical behavior. Neither the diversity nor the general properties of the proteins can be adequately explained by differ-

¹¹ Viscosity, like osmotic pressure and a large number of other properties of proteins, passes through a minimum in the neighborhood of the isoelectric point. The present discussion is limited, for lack of space, to the consideration of proteins as dipolar ions. Studies at reactions far from the isoelectric point, where proteins are present as ions, are therefore not discussed here.

ences in the size or shape of their molecules. Thus hydrocarbon molecules of the same size and shape as proteins would presumably be completely soluble in non-polar solvents and relatively unreactive to electrolytes. It is precisely because proteins are possessed of varying numbers of polar groups, distributed on their vast surfaces, that they are polar molecules, insoluble in non-polar solvents but for the most part extremely soluble in water or electrolyte solutions.

There have been great advances in recent years in estimating the amino acid compositions of proteins. The relation between cystine, methionine, and the various forms of sulfur in the protein molecule have been considered in tables 1 and 2 in connection with estimates of the minimum molecular weights of proteins. Tryptophane and tyrosine have also proved invaluable in estimating minimal molecular weights, and these amino acids can be estimated with some accuracy by colorimetric procedures (52, 53). The latter amino acid also appears to be of significance for immunity and for the combination of base by proteins as alkaline reactions.

The analytical methods for the other amino acids of importance in combining acids and bases are also being constantly improved; those for the basic amino acids among others by Vickery (149), Block (16), and Bergmann (14, 15), and those for the dicarboxylic amino acids by Foreman (54), Dakin (36), Breese Jones (70), Calverley (22), and Chibnall (24). It is not within the province of the present communication to consider these methods, but it should be pointed out that every improvement in the estimation of the amino acid composition of proteins has rendered it more certain (32) that it is by virtue of the dissociated carboxyl, sulphydryl, phenolic hydroxyl, and the positively charged imidazole, guanidine, and ammonium groups that proteins can exist as cations, as anions, or at their isoelectric points as dipolar ions. In large part the diversity in gross behavior between molecules of the same size and shape may be ascribed to differences in the number and in the distribution of these groups.

Electrophoretic mobility, electromotive force measurements, and the number of charged groups

Excepting at their isoelectric points proteins have a net charge and form salts with either acids or bases. According to the theory of dipolar ions the net charge arises at reactions alkaline to the isoelectric point from the loss of a proton from the basic amino acids, leaving the molecules negatively charged, or from the feeble acid dissociation of phenolic hydroxyl or sulphydryl groups. The regions in which the various free groups of proteins and peptides dissociate and the present state of knowledge regarding amphoteric properties have recently been discussed elsewhere (23, 31). Except for proteins with alkaline isoelectric points, the negative charges of

electrically neutral proteins arise completely, as far as we know, from the dissociation of carboxyl groups.

At reactions acid to the isoelectric point carboxyl groups become undisassociated, lose their negative charge, and leave the protein with a net positive charge upon amino, imidazole, and guanidine groups. The number of these is exactly equal, as far as present measurements reveal, to the acid-combining capacity of the proteins. Thus analytical and physico-chemical methods combine to give us accurate information regarding the maximum number of positive charges which a protein can bear as a cation or a dipolar ion.

Electrometric titration curves of proteins reveal differences ascribable to the differences in the number and dissociation of the free groups. Despite the general similarity of such curves, those for most proteins (27, 31), even for closely related proteins such as the serum globulins (58), reveal definite differences in the number or strength of dissociating groups.

Such differences in amphoteric properties must lead to differences in electrophoretic mobility, as Sir William Hardy clearly envisaged in his classical studies at the turn of the century. The calculation of net charge from measurements of electrophoretic mobility is far more complicated than from electromotive force measurements because, whereas the latter can be carried out in almost salt-free solution, the boundary conditions in mobility experiments demand the presence of adequate concentrations of buffer. Although theoretical equations for mobility of simpler ions have been developed by Onsager (110), MacInnes (88), and Sheldovsky (125), their application to proteins is not yet complete and extrapolation to salt-free conditions not yet satisfactory (see 38, 129).

Despite this temporary theoretical limitation, electrophoretic mobility measurements are extremely sensitive indications of differences in the charged condition of different proteins. Thus Tiselius, in his recent studies with an apparatus much improved over those previously available, has shown differences in the mobility of the various serum globulin fractions. Moreover, the optical systems that he has developed render possible the following of more than one boundary and can thus be employed not only to prove whether proteins of one or more valence type are present but also, in the latter case, to effect their separation (86, 141).

Whereas the basic amino acids may be present to the same extent in the various serum globulin fractions, both electromotive force measurements (58) and measurements of electrophoretic mobility (142) reveal differences in dissociation, leading to differences in isoelectric point. For these proteins we thus have a method of distinguishing between molecules presumably of the same size (91, 101), although perhaps of somewhat different shape (45). Other examples might be cited. In the case of the

crystalline serum albumins, however, no differences in molecular weight for different fractions have been shown, nor, as judged by viscosity and dielectric constant dispersion curves, in the shape of the molecules. Further, no significant difference in titration curves has been observed since the earliest studies of this kind by d'Agostino and Quagliariello (5, 27, 31). Moreover, the most recent electrophoretic mobility studies upon serum albumins (142) also reveal no difference between different fractions, although it has been known since the classical studies of Sörensen that the solubilities in salt solutions of the various crystalline serum albumins are quite different. In these proteins present methods show no differences in size, shape, or net charge, although gross differences in solubility may none the less be observed.

Dielectric constant increment, dipole moment, and the distribution of electrically charged groups

Not only the number but also the arrangement in space of the charged groups determines behavior. Thus for a tetrapole, like cystine, which might be considered as made up of two dipoles of moments equal to glycine, that is, of 15×10^{-18} electrostatic units, the moment of the whole molecule might be equal to 30×10^{-18} electrostatic units or to zero, depending on whether the dipoles were parallel or antiparallel, or had an intermediate position (34, 75, 93). In the case of an isoelectric molecule like egg albumin with twenty-eight positively and twenty-eight negatively charged groups (31), or of hemoglobin with at least seventy-five positively and negatively charged groups (32), an extremely large number of arrangements of these groups is conceivable, and each arrangement might well lead to a molecule of quite different properties.

The estimation of the dipole moments of amino acids, peptides, and proteins from measurements of dielectric constant have been considered elsewhere (30, 31, 109, 162). Studies upon egg albumin have yielded a moment of approximately 180×10^{-18} electrostatic units. The dielectric constant increments of the various egg albumin preparations that have thus far been studied by Oncley are in good agreement with each other and with the previously published studies of Shutt (126). Similarly the measurements of Oncley (107) upon horse hemoglobin are all consistent with each other and with those of Errera (43). The studies made upon horse serum albumin by Ferry and Oncley (49) reveal differences in moment corresponding to differences in solubility of the various crystalline fractions thus far investigated, ranging from 270 to 510×10^{-18} electrostatic units for fractions of low and high solubilities, respectively, in ammonium sulfate solutions. For these proteins, therefore, differences in behavior which could be ascribed to neither size, shape, nor charge can thus be ascribed to the distribution of the charge on the surface of the vast protein molecule.

Solubility and other measurements of activity coefficients and the electric moments of proteins

The earliest methods of separating proteins from each other depended upon differences in their solubilities in different solvents. Upon these differences the purification of proteins often depends, as well as the crystallization of many of them, and a classification has developed and been adopted which defines proteins as albumins, globulin, prolamines, and glutelins, depending on whether they are soluble in water, salt solutions, alcohol-water mixtures, or only in acid and alkaline solutions.

Solubility depends not only upon the forces between solvent and solute but also upon those between solute molecules in the solid state, that is, upon crystal lattice energies. The crystal lattice energies of proteins and even of amino acids and peptides are difficult to measure, since these substances cannot be sublimed or even fused. Change of solubility with change of solvent should, however, be independent of crystal lattice energy and yield activity coefficients, provided the protein is present in the same solid state in equilibrium with the various solvents investigated.

Where it has been possible to measure activity coefficients of amino acids, peptides, or proteins by means of freezing point (121), vapor pressure (128), or electromotive force measurements (71), the results, as a first approximation, have been satisfactorily in agreement with those derived from solubility measurements, whether the nitrogenous component was present as saturating body and the interaction investigated with solvent molecules (35, 47, 48), or the nitrogenous components were added to insoluble salts (46). Such studies are now too numerous to be considered here,¹⁸ but a few of the pertinent generalizations that have been deduced may be cited. (1) In regions of sufficiently low dielectric constant, interaction between ions and dipolar ions depends largely upon the electric moment and also to a small extent upon the size of the molecule. (2) Under these circumstances the solubility of elongated amino acids and peptides has been investigated, and the ratio of the logarithm of activity coefficient to ionic concentration has been shown to increase almost proportionately with the electric moment. (3) The volume of the molecule diminishes the effect due to its electric moment the more the larger the volume of the solute and the greater the dielectric constant of the solution. Thus the greater the molecular weight in comparison with the moment of the molecule the less will mutual solvent action occur, and the more will mutual precipitating action occur. The "salting-out" of proteins is the

¹⁸ Influence of dipolar ions on the solubility of each other (34) is considered in more detail in the Symposium on Intermolecular Action, held by the Division of Physical and Inorganic Chemistry of the American Chemical Society at Providence, Rhode Island, December, 1938.

best known example of this phenomenon. (4) These interactions would appear both from experience and from theoretical calculations (74, 75) to depend not only upon the sizes but also upon the shapes of the molecules and to depend not only on the dipole moment of the molecule but also upon the number of the charged groups, even where these are symmetrically arranged. Thus the dipole moment does not account completely for molecular interactions between dipolar ions (48), but higher electric moments must presumably also be considered.

IV. DISCUSSION

The adequate development of a theory of protein chemistry demands a formulation of the parameters in terms of which behavior may be explained. In the foregoing we have attempted to discuss certain of these parameters. Certainly the size of the protein molecule must be considered. Knowledge of the size of a protein is not, however, sufficient for its characterization. Egg albumin, pepsin, lactoglobulin, and zein all have nearly the same molecular weight. Yet the first is classified as a water-soluble albumin, the second and third as globulins (insoluble in water but more soluble in salt solutions), and zein as a prolamine insoluble in water but soluble in alcohol-water mixtures.

The shape of a protein is not sufficient for its characterization. Fibrinogen, among the blood proteins, is rod-shaped and so is tobacco mosaic virus. Not all blood proteins nor all viruses are elongated molecules, however. Among the former, hemoglobin is nearly spherical, and there is at least one virus, the bushy stunt virus of tomato (13, 92), which is reported to be nearly spherical.

The isoelectric point of a protein is not sufficient for its characterization. Egg and serum albumins, casein, and gelatin have nearly the same isoelectric point. Although the first two have approximately equal numbers of free acid and basic groups, casein has many more acid than basic, and gelatin many more basic than acid groups.

The dielectric constant increment of a protein is not sufficient for its characterization. Carboxyhemoglobin and the most soluble crystalline fraction of horse serum albumin increase the dielectric constant of water by nearly the same amount, but have quite different relaxation times.

The relaxation time of a protein is not sufficient for its characterization. The different fractions of crystalline horse serum albumin have the same relaxation time, though the dielectric increment of the more soluble is roughly three times that of the less soluble. Moreover, the relaxation time—and also the viscosity—of carboxyhemoglobin is smaller than that of the serum albumins, although all have nearly the same molecular weight.

Whereas no one of these properties will suffice for characterization, these

and other measurements, taken together, yield quantitative information in terms of which we may develop still further both the theory and the practice of protein chemistry.

REFERENCES

- (1) ADAIR: Proc. Cambridge Phil. Soc. **1**, 75 (1924); Skand. Arch. Physiol. **49**, 76 (1926).
- (2) ADAIR: Proc. Roy. Soc. (London) **A120**, 573 (1928); J. Am. Chem. Soc. **51**, 696 (1929).
- (3) ADAIR AND ADAIR: Proc. Roy. Soc. (London) **B120**, 422 (1936).
- (4) ADAIR AND ROBINSON: Biochem. J. **24**, 1864 (1930).
- (5) D'AGOSTINO AND QUAGLIARIBELLO: Nernst Festschrift, Halle, p. 27 (1912).
- (6) ARRHENIUS: J. Chem. Phys. **5**, 63 (1937).
- (7) ARRHENIUS: Physik. Z. **39**, 559 (1938).
- (8) BAERNSTEIN: J. Biol. Chem. **89**, 125 (1930).
- (9) BAERNSTEIN: J. Biol. Chem. **97**, 663 (1932); **106**, 451 (1934); **115**, 25 (1936).
- (10) BAERNSTEIN: J. Biol. Chem. **97**, 669 (1932); **115**, 33 (1936).
- (11) BAILEY: Biochem. J. **31**, 1396 (1937).
- (12) BAILEY: Biochem. J. **31**, 1406 (1937).
- (13) BAWDEN AND PIRIE: Nature **141**, 513 (1938); British J. Exptl. Path. **19**, 251 (1938).
- (14) BERGMANN: Chem. Rev. **22**, 423 (1938).
- (15) BERGMANN AND NIEMAN: J. Biol. Chem. **122**, 577 (1938).
- (16) BLOCK: The Determination of Amino Acids, Minneapolis (1938).
- (17) BLUMENTHAL AND CLARK: J. Biol. Chem. **110**, 343 (1935).
- (18) BURK: J. Biol. Chem. **98**, 353 (1932).
- (19) BURK: J. Biol. Chem. **120**, 63 (1937); **121**, 373 (1937); **124**, 49 (1938).
- (20) BURK AND GREENBERG: J. Biol. Chem. **87**, 197 (1930).
- (21) BOEHM AND SIGNER: Helv. Chim. Acta **14**, 1370 (1931); Klin. Wochschr. **11**, 599 (1932).
- (22) CALVERY: J. Biol. Chem. **94**, 613 (1931-32).
- (23) CANNAN: Cold Spring Harbor Symposia Quant. Biol. **6**, 1 (1938).
- (24) CHIBNALL: Personal communication.
- (25) CHICK: Biochem. J. **8**, 261 (1914).
- (26) CHICK AND LURZYNNSKA: Biochem. J. **8**, 59 (1914).
- (27) COHN: Physiol. Rev. **5**, 349 (1925).
- (28) COHN: Ergeb. Physiol. biol. Chem. exptl. Pharmakol. **33**, 781 (1933).
- (29) COHN: Naturwissenschaften **20**, 663 (1932).
- (30) COHN: Ann. Rev. Biochem. **4**, 93 (1935).
- (31) COHN: Cold Spring Harbor Symposia Quant. Biol. **6**, 8 (1938).
- (32) COHN, GREEN, AND BLANCHARD: J. Am. Chem. Soc. **59**, 509 (1937).
- (33) COHN, HENDRY, AND PRENTISS: J. Biol. Chem. **63**, 721 (1925).
- (34) COHN, McMEEKIN, FERRY, AND BLANCHARD: J. Phys. Chem. **43**, 169 (1939).
- (35) COHN, McMEEKIN, GREENSTEIN, AND WEARE: J. Am. Chem. Soc. **58**, 2365 (1936).
- (36) DAKIN: Biochem. J. **12**, 290 (1918).
- (37) DANIEL AND COHN: J. Am. Chem. Soc. **58**, 415 (1936).
- (38) DAVIS AND COHN: Unpublished work.
- (39) DEBYE: Polar Molecules. The Chemical Catalog Company, Inc., New York (1929).

- (40) EHRLICH: *Z. Ver. deutsch. Zucker-Ind.* **53**, 809 (1903).
(41) ELLIOTT AND WILLIAMS: *J. Am. Chem. Soc.* **61**, 718 (1939).
(42) ERIKSSON-QUENSEL AND SVEDEBERG: *Biol. Bull.* **71**, 498 (1936).
(43) ERRERA: *J. chim. phys.* **29**, 577 (1932).
(44) EYRING: *Phys. Rev. [2]* **39**, 746 (1932).
(45) FAHEY AND GREEN: *J. Am. Chem. Soc.* **60**, 3039 (1938).
(46) FAILEY: *J. Am. Chem. Soc.* **54**, 576, 2367 (1932); **55**, 4374 (1933).
(47) FERRY, COHN, AND NEWMAN: *J. Am. Chem. Soc.* **58**, 2370 (1936).
(48) FERRY, COHN, AND NEWMAN: *J. Am. Chem. Soc.* **60**, 1480 (1938).
(49) FERRY AND ONCLEY: *J. Am. Chem. Soc.* **60**, 1128 (1938).
(50) FERRY AND ONCLEY: Unpublished work.
(51) FISCHER: *Z. physiol. Chem.* **33**, 151 (1901); *Ber.* **35**, 2660 (1902).
(52) FOLIN AND CIOCALTEU: *J. Biol. Chem.* **73**, 627 (1927).
(53) FOLIN AND MARENZI: *J. Biol. Chem.* **83**, 89 (1929).
(54) FOREMAN: *Biochem. J.* **13**, 378 (1919).
(55) FRIEND, FERRY, AND ONCLEY: *J. Biol. Chem.* **125**, Proc. 39 (1938).
(56) FÜRTH AND LIEBEN: *Biochem. Z.* **109**, 124 (1920).
(57) GANS: *Ann. Physik* **86**, 628 (1928).
(58) GREEN: *J. Am. Chem. Soc.* **60**, 1108 (1938).
(59) GREENSTEIN: *J. Biol. Chem.* **125**, 501 (1938).
(60) GUTH: *Kolloid-Z.* **74**, 147 (1936).
(61) GUTH AND MARK: *Ergeb. exakt. Naturw.* **12**, 115 (1933).
(62) HAUGAARD AND JOHNSON: *Compt. rend. trav. lab. Carlsberg* **18^a**, 96 (1930).
(63) HERZOG, ILLIG, AND KUDAR: *Z. physik. Chem.* **A167**, 329 (1934).
(64) HEWITT: *Biochem. J.* **30**, 2229 (1936); **31**, 380 (1937).
(65) HOPKINS: *Biochem. J.* **15**, 286 (1921); *Nature* **126**, 328, 383 (1930).
(66) HUANG AND WU: *Chinese J. Physiol.* **4**, 221 (1930).
(67) HÜFNER AND GANSSEER: *Arch. Physiol.* p. 209 (1907).
(68) HUGGINS: *J. Phys. Chem.* **42**, 911 (1938).
(69) HUNTER AND BOBOOK: *J. Biol. Chem.* **57**, 507 (1923).
(70) JONES AND WILSON: *Cereal Chem.* **5**, 473 (1928).
(71) JOSEPH: *J. Biol. Chem.* **111**, 479, 489 (1935).
(72) KASSEL AND BRAND: *Proc. Soc. Exptl. Biol. Med.* **35**, 444 (1936).
(73) KASSEL AND BRAND: *J. Biol. Chem.* **125**, 435 (1938).
(74) KIRKWOOD: *J. Chem. Phys.* **2**, 351 (1934).
(75) KIRKWOOD: *Chem. Rev.* **24**, 233 (1939).
(76) KOSSEL: *Z. physiol. Chem.* **22**, 178 (1896-7).
(77) KRÜGER: *Arch. ges. Physiol. (Pflügers)* **43**, 244 (1888).
(78) KUHN: *Z. physik. Chem.* **A161**, 1 (1932).
(79) KUHN: *Kolloid.-Z.* **68**, 2 (1934); *Z. physik. Chem.* **A175**, 1 (1935).
(80) KUHN AND DESNOUELLE: *Z. physiol. Chem.* **251**, 14 (1938).
(81) KÜLZ: *Z. Biol.* **27**, 415 (1890).
(82) LAMM AND POLSON: *Biochem. J.* **30**, 528 (1936).
(83) LAUFFER AND STANLEY: *J. Biol. Chem.* **123**, 507 (1938).
(84) LINDEMSTRØM-LANG: *Compt. rend. trav. lab. Carlsberg* **17**, No. 9 (1929).
(85) LOEB: *Proteins and the Theory of Colloidal Behavior*, 2nd edition. McGraw-Hill Book Company, New York (1924).
(86) LONGSWORTH AND MACINNES: *Chem. Rev.* **24**, 271 (1939).
(87) LOONEY: *J. Biol. Chem.* **69**, 519 (1926).
(88) MACINNES, SHEDLOVSKY, AND LONGSWORTH: *Chem. Rev.* **13**, 29 (1933).

- (89) MARRACK AND HEWITT: Biochem. J. **23**, 1079 (1929).
(90) MCCOY, MEYER, AND ROSE: J. Biol. Chem. **112**, 283 (1935-36).
(91) McFARLANE: Biochem. J. **29**, 407, 660, 1209 (1935).
(92) McFARLANE AND KEKWICK: Biochem. J. **32**, 1607 (1938).
(93) McMEEKIN, COHN, AND BLANCHARD: J. Am. Chem. Soc. **59**, 2717 (1937).
(94) MEHL: Cold Spring Harbor Symposia Quant. Biol. **6**, 218 (1938).
(95) MILLER AND DU VIGNEAUD: J. Biol. Chem. **118**, 101 (1937).
(96) MIRSKY AND ANSON: J. Gen. Physiol. **18**, 307 (1935).
(97) MORNER: Z. physiol. Chem. **28**, 595 (1899).
(98) MUELLER: J. Biol. Chem. **56**, 157 (1923).
(99) MULDER: Ann. Chemie **28**, 73 (1838).
(100) VON MURALT AND EDSELL: J. Biol. Chem. **89**, 315, 351 (1930).
(101) VON MUTZENBECHER: Biochem. Z. **266**, 226 (1933).
(102) VON MUTZENBECHER: Biochem. Z. **266**, 250 (1933).
(103) VON MUTZENBECHER AND SVEDBERG: Naturwissenschaften **21**, 331 (1933).
(104) NEURATH: Cold Spring Harbor Symposia Quant. Biol. **6**, 196 (1938).
(105) NICHOLS: J. Am. Chem. Soc. **52**, 5176 (1930).
(106) NORTHROP AND ANSON: J. Gen. Physiol. **12**, 543 (1929); **20**, 575 (1936-37).
(107) ONCLEY: J. Am. Chem. Soc. **60**, 1115 (1938).
(108) ONCLEY: Unpublished work.
(109) ONCLEY, FERRY, AND SHACK: Cold Spring Harbor Symposia Quant. Biol. **6**, 21 (1938).
(110) ONSAGER: Physik. Z. **27**, 388 (1926); **28**, 277 (1927).
(111) OSBORNE: J. Am. Chem. Soc. **24**, 140 (1902); *The Vegetable Proteins*, Longmans, Green and Co., London and New York (1909).
(112) PALMER: J. Biol. Chem. **104**, 359 (1934).
(113) PEDERSEN: Nature **138**, 363 (1936).
(114) PEDERSEN: Biochem. J. **30**, 945 (1936).
(115) PEDERSEN: Compt. rend. trav. lab. Carlsberg **22**, 427 (1938).
(116) PEDERSEN: Unpublished work.
(117) PERRIN: J. phys. radium **7**, 1 (1936).
(118) POLSON: Nature **137**, 740 (1936).
(119) ROCHE AND MARQUET: Compt. rend. trav. lab. Carlsberg **118**, 898 (1935).
(120) ROSE: J. Biol. Chem. **112**, 283 (1935-36); Physiol. Rev. **18**, 109 (1938).
(121) SCATCHARD AND PRENTISS: J. Am. Chem. Soc. **56**, 2314 (1934).
(122) SCHULTZ: Z. physiol. Chem. **25**, 16 (1898).
(123) SCOTT AND FISHER: Biochem. J. **29**, 1048 (1935).
(124) SHACK: Unpublished work.
(125) SHEDLOVSKY: J. Am. Chem. Soc. **54**, 1405 (1932); J. Franklin Inst. **225**, 739 (1938).
(126) SHUTT: Trans. Faraday Soc. **30**, 893 (1934).
(127) SJOGREN AND SVEDBERG: J. Am. Chem. Soc. **53**, 2657 (1931).
(128) SMITH AND SMITH: J. Biol. Chem. **117**, 209 (1937); **121**, 607 (1937).
(129) SMITH: J. Biol. Chem. **108**, 187 (1935); **113**, 473 (1936).
(130) SØRENSEN: Compt. rend. trav. lab. Carlsberg **12**, 282 (1917).
(131) SØRENSEN: Compt. rend. trav. lab. Carlsberg **18**, No. 5, 1 (1930).
(132) STAUDINGER: Die hochpolymeren organischen Verbindungen. Julius Springer, Berlin (1932).
(133) STEINHARDT: J. Biol. Chem. **123**, 543 (1938).
(134) SULLIVAN AND HESS: Pub. Health Repts., Suppl. No. 86 (1930).

- (135) SVEDBERG: Chem. Rev. **14**, 1 (1934); **20**, 81 (1937); Kolloid-Z. **85**, 119 (1938).
- (136) SVEDBERG, CARPENTER, AND CARPENTER: J. Am. Chem. Soc. **52**, 241 (1930).
- (137) SVEDBERG, CARPENTER, AND CARPENTER: J. Am. Chem. Soc. **52**, 701 (1930).
- (138) SVEDBERG AND FAHRAEUS: J. Am. Chem. Soc. **48**, 430 (1926).
- (139) SVEDBERG AND HEDENIUS: Biol. Bull. **64**, 191 (1934).
- (140) SVEDBERG AND NICHOLS: J. Am. Chem. Soc. **48**, 308 (1926).
- (141) TISELIUS: Dissertation, Upsala, 1930.
- (142) TISELIUS: Biochem. J. **31**, 1464 (1937).
- (143) TISELIUS AND GROSS: Kolloid-Z. **66**, 11 (1934).
- (144) TOENNIS: Growth **1**, 337 (1937).
- (145) VAN SLYKE: Handb. d. biol. Arbeitsmeth., Abt. I, Teil 7, 51 (1923).
- (146) VAN SLYKE AND BAKER: J. Biol. Chem. **35**, 127 (1918).
- (147) VAN SLYKE, HILLER, DILLON, AND MACFADYEN: Proc. Soc. Exptl. Biol. Med. **38**, 548 (1938).
- (148) VICKERY: Compt. rend. trav. lab. Carlsberg **22**, 519 (1938)
- (149) VICKERY AND LEAVENWORTH: J. Biol. Chem. **79**, 377 (1928).
- (150) VICKERY AND WHITE: J. Biol. Chem. **99**, 701 (1933)
- (151) DU VIGNEAUD, JENSEN, AND WINTERSTEINER: J. Pharmacol. **32**, 367 (1928).
- (152) WATSON, ARRHENIUS, AND WILLIAMS: Nature **137**, 322 (1936).
- (153) WEBER: Ergeb. Physiol. biol. Chem. exptl. Pharmakol. **36**, 109 (1934).
- (154) WEBER AND STOVER: Biochem. Z. **269**, 269 (1933).
- (155) WILLIAMS: Trans. Faraday Soc. **30**, 723 (1934).
- (156) WILLIAMS AND ONCLEY: Physics **3**, 314 (1932)
- (157) WILLIAMS AND WATSON: Nature **139**, 506 (1937); Cold Spring Harbor Symposia Quant. Biol. **6**, 208 (1938).
- (158) WOLLASTON: Phil. Trans. Roy. Soc., p. 223 (1810).
- (159) WRINCH: Proc. Roy. Soc. (London) **A160**, 59; **A161**, 505 (1937); Trans. Faraday Soc. **33**, 1368 (1937).
- (160) WU AND YANG: Chinese J. Physiol. **6**, 51 (1932).
- (161) WYMAN: J. Biol. Chem. **90**, 443 (1931).
- (162) WYMAN: Chem. Rev. **19**, 213 (1936); J. Am. Chem. Soc. **60**, 328 (1938).
- (163) WYMAN: J. Phys. Chem. **43**, 143 (1939).
- (164) ZAHND AND CLARKE: J. Biol. Chem. **102**, 171 (1933).

THEORETICAL STUDIES UPON DIPOLAR IONS¹

J. G. KIRKWOOD

Department of Chemistry, Cornell University, Ithaca, New York

Received January 30, 1939

I

The influence of electrolytes on the behavior of the amino acids and proteins is perhaps one of the most important topics in the physical chemistry of these substances. It has long been known that there is a strong thermodynamic interaction between electrolytes and the amino acids and proteins. In recent years the accurate measurements of Cohn and his coworkers (1, 2, 3, 4, 5) have provided a large body of thermodynamic data relating to this question.

With the accumulation of evidence in favor of the dipolar ion or zwitter-ion structure of the amino acids and proteins in the isoelectric condition in solvents of high dielectric constant, it has become apparent that thermodynamic interaction between electrolytes and these substances can, in large measure, be attributed to the strong electrostatic intermolecular forces between dipolar ions and real ions in solution. A dipolar ion, while bearing no net charge, is characterized by electric multipole moments of large magnitude. For example, the glycine dipolar ion, $\text{NH}_3^+ \text{CH}_2\text{COO}^-$, possesses a dipole moment of about 15 Debye units, about ten times that of an ordinary polar molecule. Thus a dipolar ion is in a sense a superpolar molecule, surrounded by an intense electrostatic field.

An extension of the Debye-Hückel theory describing the electrostatic interaction of dipolar ions and real ions in solution was developed by Scatchard and the writer (10) and later elaborated in some detail by the writer (8) for dipolar ions of spherical shape. It is the purpose of the present article to review the possible applications of the theory and to extend it to dipolar ions of elongated shape. In this extension of the theory we shall confine ourselves to limiting laws approached at high

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

dilution, making use of a method similar to that developed by Fuoss (6) for the study of dipole-dipole interaction, rather than the somewhat more complicated method based upon the Debye-Hückel theory.

The theory may be employed in two ways. By means of it the influence of electrolytes upon the chemical potential of a dipolar ion of known electrical structure may be predicted. On the other hand, it may be used to determine the structure, for example, the distance between the charged groups and the dipole moments, of simple dipolar ions from data relating to the influence of electrolytes on their chemical potentials. For the latter purpose we shall make extensive use of the measurements of Cohn and his coworkers.

In this article we shall limit the discussion to the aliphatic amino acids and their peptides, since the proteins present certain ambiguities of interpretation with which we cannot concern ourselves here. Thus, so many parameters are required to specify the charge distribution of protein dipolar ions that they cannot be uniquely determined from the influence of electrolyte upon chemical potential. On the other hand, too little is known at present about the details of protein structure to permit a probable assignment of configuration to the dipolar ionic charges from which to predict the interaction with electrolytes. Nevertheless, as Cohn has shown, certain progress can be made along the latter lines, and observed electrolyte effects can, in the case of egg albumin and hemoglobin (5), be reproduced by hypothetical though not unique assignments of charge configuration.

II

We shall be interested in the properties of a solution containing a dipolar ion component i at a molar concentration, c_i , and ν ionic components at molar concentrations, $c_1 \dots c_\nu$, in a solvent of dielectric constant D . The pertinent thermodynamic properties of the solution may be derived from the chemical potentials of the several components. We shall be particularly concerned with the chemical potential of the dipolar ion species, which may be written in conventional form,

$$\begin{aligned}\mu_i &= RT \log \gamma_i c_i + \mu_i^0(T, p) \\ \mu_i^0 &= \lim_{c_1 \dots c_\nu \rightarrow 0} [\mu_i - RT \log c_i]\end{aligned}\quad (1)$$

where γ_i , the activity coefficient, is defined by this equation.

We shall make the simplifying assumption that the deviation of the solution from ideal behavior is due to electrostatic intermolecular forces

alone. If the solvent is idealized as a structureless dielectric continuum, the logarithm of the activity coefficient, γ_i , may be written as follows (8),

$$\log \gamma_i = \sum_{k=1}^{n+1} B_{ik} c_k$$

$$B_{ik} = \frac{N\beta}{1000} \int_0^v \int_0^1 V_{ik} e^{-\beta W_{ik}(\lambda)} d\lambda dv \quad (2)$$

$$\beta = 1/kT$$

where the sum extends over all $n + 1$ solute components. V_{ik} is the electrostatic work required to bring the pair of molecules or ions, i and k , from infinite separation to the given configuration in the pure solvent, and $W_{ik}(\lambda)$ is the average work (potential of average force) expended in the same process in the actual solution, all charges, $e_1 \dots e_n$, of molecule i having a fraction, λ , of their full values. The integration extends over all values of the relative coördinates of the pair of molecules in the volume, v , of the solution and outside a region, ω , of molecular dimensions determined by the size and shape of the two molecules, into which intermolecular repulsion at short range prevents penetration.²

According to the Debye-Hückel theory the potentials of average force, W_{ik} , necessary for the evaluation of the coefficients, B_{ik} , satisfy the following relation,

$$W_{ik}(\lambda) = z_k e \psi_i(\lambda, r_k) \quad (3)$$

where ψ_i is the average electrostatic potential at the point r_k from an origin in molecule i and $z_k e$ is the charge of ion k . The potential, ψ_i , satisfies the Poisson-Boltzmann equation of the Debye-Hückel theory, which in its linear approximation is

$$\nabla^2 \psi_i - \kappa^2 \psi_i = 0 .$$

$$\kappa^2 = \frac{4\pi N e^2}{1000 D k T} \sum_{k=1}^n c_k z_k^2 \quad (4)$$

² Equation 2, while somewhat more convenient for our purposes, is essentially equivalent to the more usual expression, based on the Gunterberg-Müller charging process.

$$kT \log \gamma_i = \sum_{i=1}^n e_i \int_0^1 \{\psi_i(\lambda e_1, \dots, \lambda e_n) - \psi_i^0(\lambda e_1, \dots, \lambda e_n)\} d\lambda$$

where ψ_i is the average electrostatic potential at the point of location of charge, e_i , of molecule i in the actual solution and ψ_i^0 is its value at infinite dilution, all charges, $e_1 \dots e_n$, having a fraction, λ , of their full values.

outside the region of non-penetration, while in ω Laplace's equation is satisfied.

$$\nabla^2 \psi_i = 0 \quad (5)$$

Solution of equations 4 and 5, subject to the boundary conditions of electrostatics on the surface of ω , continuity of the potential and the normal component of the dielectric displacement, yields ψ_i , which may be used in equation 2 or 2a for the calculation of the activity coefficient. This method has been employed by the writer to study the influence of electrolytes on the activity of spherical dipolar ions (8).

In the present article we propose to use a somewhat simpler method, which yields only a limiting law. Only linear terms in the expansion of the logarithm of the activity coefficient in the concentrations of the various solute species appear, thus allowing the treatment of non-spherical dipolar ions and permitting the inclusion of "salting-out" forces between an ion and a dipolar ion. The latter arise from a repulsion between the ionic charge and an image distribution in the cavity of low dielectric constant created by the dipolar ion in the solvent. It is evident that

$$\lim_{c_1, \dots, c_{s+1} \rightarrow 0} W_{ik}(\lambda) = \lambda V_{ik} \quad (6)$$

If we are content with the linear terms in a power series in the ionic concentrations, we may, therefore, write equation 2 in the following form,

$$\begin{aligned} \log \gamma_i &= \sum_{k=1}^s B_{ik}^0 c_k \\ B_{ik}^0 &= \frac{N}{1000} \int_{\omega}^{\infty} (1 - e^{-\beta v \cdot k}) dv \\ B_{ik}^0 &= \lim_{c_1, \dots, c_{s+1} \rightarrow 0} B_{ik} \end{aligned} \quad (7)$$

For simplicity we suppose that the term in the sum, equation 2, arising from the mutual interaction of dipolar ions is negligible and thus limit ourselves to solutions in which the concentration of dipolar ions is itself small. If desired, the B_{ik}^0 for mutual dipolar interaction may be obtained from Fuoss' calculations for dipole molecules (6). The development (equation 7) is only possible when the integrals defining the B_{ik} exist. Although they diverge when the component i is a true ion, they exist when it is a dipolar ion, bearing no net charge, and the expansion is suitable for our purposes. In solvents of relatively high dielectric constant at ordinary temperatures, a satisfactory approximation to the B_{ik} may be

obtained by expansion of $1 - e^{-\beta V_{ik}}$ with retention of the first two terms alone.

$$B_{ik}^0 = \frac{N}{1000} \int_{\omega}^{\infty} [\beta V_{ik} - \frac{\beta^2}{2} V_{ik}^2] dv \quad (8)$$

For a dipolar ion characterized by a set of charges, $e_1 \dots e_s$, distributed in a cavity, ω_0 , of dielectric constant D_0 , the electrical work, V_{ik} , required to bring a true ion of charge, $z_k e$, from infinity to a point r_k from an origin fixed in the dipolar ion i is given by

$$V_{ik} = \frac{1}{2} \left\{ z_k e \psi_e^0(r_k) + \sum_{i=1}^s e_i \psi_i^0(r_i) \right\} \quad (9)$$

where $\psi_i^0(r_i)$ is the electrostatic potential in the interior of ω_0 at the location of the charge, e_i , and $\psi_e^0(r_k)$ is the potential at the point r_k exterior to ω_0 , when the ion and dipolar ion are fixed in the given configuration in the pure solvent. Self-energy terms in equation 9 are to be omitted, V_{ik} vanishing when the ions are at infinite separation. The potentials ψ_i^0 and ψ_e^0 satisfy Laplace's equation, $\nabla^2 \psi = 0$, both interior and exterior to the surface of ω_0 , and fulfill the usual boundary conditions everywhere on this surface,

$$\begin{aligned} \psi_i^0 &= \psi_e^0 \\ D_{0n} \cdot \nabla \psi_i^0 &= D_n \cdot \nabla \psi_e^0 \end{aligned} \quad (10)$$

where n is a unit vector normal to the surface. In addition, the potentials must have the singularities characteristic of the real charge distributions of the two ions. The details of the determination of the electrostatic potential for dipolar ions of several shapes are given in the appendix.

We first consider a spherical dipolar ion of radius b , the charge distribution of which may be characterized by a point dipole of moment, μ , located at its center. The cavity, ω_0 , is thus a sphere of radius b , and the excluded region, ω , in the integral (equation 8) is a sphere of radius a , the sum of the radii of the dipolar ion and the real ion, all real ions in the solution being spheres of the same radius. The potential energy, V_{ik} , has the following form,

$$V_{ik} = \frac{3z_k e \mu \cos \vartheta}{(2D + D_0)r^2} + \frac{z_k^2 e^2 b^4}{2\pi^4} \frac{D - D_0}{D} \sum_{n=0}^{\infty} \frac{n+1}{(n+2)D + (n+1)D_0} \left(\frac{b}{r}\right)^{2n} \quad (11)$$

where r is the distance between the centers of the ion and dipolar ion, and ϑ is the angle between the vectors, r and μ , the dipole moment. The second member of equation 11 represents a repulsion between the real

ion and an image distribution in the cavity, ω_0 , created by the dipolar ion in the solvent. A similar term, usually of much smaller magnitude, arising from the interaction of the dipolar ion and an image distribution in the ionic cavity has been suppressed.

Substitution of equation 11 in equations 7 and 8 and the neglect of small terms in D_0/D in the summation of the resulting series yields the following limiting law for the activity coefficient of the dipolar ion component of the solution,

$$\begin{aligned} \log_{10} \gamma_i &= -B_i \Gamma \\ \Gamma &= (1/2) \sum_{k=1}^{\infty} c_k z_k^2 \\ B_i &= \frac{2\pi N e^2}{2303 D k T} \left\{ \frac{3}{2} \frac{\mu^2}{D a k T} - \frac{b^3}{a} \alpha(\rho) \right\} \end{aligned} \quad (12)$$

where Γ is the ionic strength of the solution, ρ is the ratio b/a , and $\alpha(\rho)$ is a function tabulated in table 1. Insertion of numerical values for the

TABLE 1

$$\alpha(\rho) = (\frac{1}{2}\rho^4)\{(\rho^2 - 2)\log(1 + \rho) - (\rho^2 + 2)\log(1 - \rho) - 2\rho^2\}$$

| ρ | $\alpha(\rho)$ |
|--------|----------------|
| 0.0 | 1.00 |
| 0.2 | 1.01 |
| 0.4 | 1.08 |
| 0.6 | 1.21 |
| 0.8 | 1.54 |
| 0.9 | 1.96 |

universal constants and the introduction of the dielectric constant of water, D_w , yields at 25°C.

$$\begin{aligned} B_i &= B_i^{(0)} - B_i^{(1)} \\ B_i^{(0)} &= 5.48 \times 10^{-3} (D_w/D)^2 \mu^2 / a = 0.125 (D_w/D)^2 R^2 / a \quad (13) \\ B_i^{(1)} &= 4.66 \times 10^{-3} (D_w/D) (v_i/a) \alpha(\rho) \end{aligned}$$

where R is the effective dipole separation, μ/e , and v_i is equal to $4\pi N b^3/3$, the value at infinite dilution of the partial molal volume of the dipolar ion component under the idealization of the solvent as structureless continuum. The units employed in equation 13 are the Debye unit for dipole moment and the Ångström unit for the lengths R and a . The term $B_i^{(0)}$ gives rise to a decrease in the chemical potential of the dipolar ion by electrolyte ("salting-in"), while the term $B_i^{(1)}$ gives rise to an increase ("salting-out"). The latter effect was not included in the writer's previous

treatment (8), based on the Debye-Hückel theory. If the term B_i^0 is omitted in equation 12, a limiting law in exact agreement with the writer's earlier one is reached.

We next consider a dipolar ion of prolate ellipsoidal shape, characterized by a charge distribution consisting of two charges, $+e$ and $-e$, located at the foci. It is convenient to employ confocal elliptical coördinates in the treatment of this model. Thus a point situated at distances r_1 and r_2 from the respective foci of the ellipsoid and in a plane inclined at an angle, φ , to the reference plane containing the major axis is specified by the coördinates λ , η , and φ .

$$\begin{aligned}\lambda &= (r_1 + r_2)/R \\ \eta &= (r_1 - r_2)/R\end{aligned}\quad (14)$$

where R is the interfocal distance. The cavity, ω_0 , is an ellipsoid of eccentricity, ϵ , equal to $1/\lambda_0$, and the region ω , into which ions cannot penetrate is, for simplicity, assumed identical with ω_0 , although this is strictly true only for vanishing ion size. The potential energy, V_{ik} , of the dipolar ion and a real ion of charge $z_k e$ is closely approximated by the following expression,

$$V_{ik} = \frac{4z_k e^2}{DR} \frac{\eta}{\lambda^2 - \eta^2} \lambda_0^{-1} \left[\lambda_0 - \frac{(\lambda_0^2 - 1)}{2} \log \frac{\lambda_0 + 1}{\lambda_0 - 1} \right] \quad (15)$$

Equation 15 is exact for the two limiting cases, zero and unit eccentricity, and only slightly inaccurate, owing to the approximate summation of an infinite series, for intermediate eccentricities. The term arising from the interaction of the ion with its image distribution in the dipolar ion cavity, included for the sphere, has been omitted in equation 15, since it introduces undue complication, which we postpone for later, in the evaluation of the integral (equation 8). Introduction of the coördinates λ and η as variables of integration in the latter equation leads to the following expression

$$B_{ik}^0 = - \frac{\pi N \beta^2 R^3}{1000} \int_{\lambda_0}^{\infty} \int_{-1}^{+1} (\lambda^2 - \eta^2) V_{ik}^2 d\eta d\lambda \quad (16)$$

the term linear in V_{ik} vanishing since it is an odd function of η . Equation 16, together with equations 7 and 15, yields for the logarithm of the activity coefficient of the dipolar ion,

$$\begin{aligned}\log_{10} \gamma_i &= - B_i \Gamma \\ B_i &= \frac{2\pi N e^4 g(\lambda_0)}{2303(DkT)^2} R \\ g(\lambda_0) &= \lambda_0^{-2} \left[\lambda_0 - \frac{(\lambda_0^2 - 1)}{2} \log \frac{\lambda_0 + 1}{\lambda_0 - 1} \right]^{-1}\end{aligned}\quad (17)$$

Insertion of numerical values for the constants gives, at 25°C.,

$$B_s = 0.167(D_w/D)^2 g(\lambda_0) R \quad (18)$$

A tabulation of g as a function of the eccentricity of the ellipsoidal cavity is given in table 2. We remark that for constant eccentricity B_s is proportional to the first power of the distance, R , between the charges of the dipolar ion. For elongated ellipsoids in the neighborhood of unit eccentricity, the function g is approximately unity. For ellipsoids nearly spherical, g may be expanded in a power series in R/a , where a is the shortest distance of either focus to the surface, equal to the radius of the sphere at zero eccentricity. The initial term of the series, $3R/4a$, when substituted in equation 18 yields an expression identical with $B_s^{(0)}$ (equation 13) for a dipole at the center of a sphere. No counterpart of $B_s^{(1)}$ is obtained, since we have neglected the "salting-out" influence of image

TABLE 2
g as a function of the eccentricity of the ellipsoidal cavity

| ϵ | $g(\epsilon)$ | $\epsilon(1 - \epsilon^2)^{-1/2}g(\epsilon)$ |
|------------|---------------|----------------------------------------------|
| 0.00 | 0.00 | 0.00 |
| 0.20 | 0.30 | 0.061 |
| 0.33 | 0.49 | 0.17 |
| 0.50 | 0.71 | 0.39 |
| 0.60 | 0.83 | 0.58 |
| 0.70 | 0.94 | 0.82 |
| 0.80 | 1.01 | 1.14 |
| 1.00 | 1.00 | |

forces in the present case. While we shall not consider the problem in detail in the present article, it seems reasonable to suppose that the magnitude of B_s for an ellipsoid can be roughly estimated by an analog of the spherical formula, equation 13, in which v_s/l appears instead of v_s/a , l being a length intermediate between the semi-major and semi-minor axes of the ellipsoid.

We shall also discuss a second model in which the dipolar ion is ellipsoidal in shape but characterized by a charge distribution consisting of a point dipole of moment, μ , situated at one of the foci and parallel to the major axis. The calculations proceed in the same manner as for the first ellipsoidal model except that the potential, $V_{s,k}$, is of the form,

$$V_{s,k} = \frac{4\pi z_k e}{DR^2} \frac{\lambda\eta}{(\lambda + \eta)^3} \lambda_0^{-1} \left[\lambda_0 - \frac{(\lambda_0^2 - 1)}{2} \log \frac{\lambda_0 + 1}{\lambda_0 - 1} \right]^{-1} \quad (19)$$

The logarithm of the activity coefficient of the dipolar ion is finally calculated to be

$$\log_{10} \gamma_i = - B_i T$$

$$B_i = \frac{3\pi Ne^2 u(\lambda_0) \mu^2}{2303(DkT)^2 R} \quad (20)$$

$$u(\lambda_0) = \frac{8}{9} \cdot \frac{1 + 3(\lambda_0^{\frac{2}{3}} - 1)^{-1} + 4(\lambda_0^{\frac{2}{3}} - 1)^{-2}}{\lambda_0(\lambda_0^{\frac{2}{3}} - 1) \left[\lambda_0 - \frac{(\lambda_0^{\frac{2}{3}} - 1)}{2} \log \frac{\lambda_0 + 1}{\lambda_0 - 1} \right]}$$

With numerical values for the constants at 25°C., B_i is given by the following equation,

$$B_i = 5.48 \times 10^{-3} (D_w/D)^2 u(\lambda_0) \mu^2 / R \quad (21)$$

in which the Debye and the Ångström are the units employed. The function $u(\lambda_0)$ is listed for several eccentricities in table 3. In the limit

TABLE 3
Values of the function $u(\lambda_0)$

| ϵ | $u(\epsilon)$ | $\epsilon(1 - \epsilon^2)^{-1/2} [u(\epsilon)]^{-1}$ |
|------------|---------------|------------------------------------------------------|
| 0.00 | 0.00 | 0.50 |
| 0.20 | 0.47 | 0.43 |
| 0.33 | 1.03 | 0.33 |
| 0.50 | 2.92 | 0.19 |
| 0.60 | 6.31 | 0.11 |
| 0.70 | 16.5 | 0.053 |
| 0.80 | 81.5 | 0.014 |

of zero eccentricity of the ellipsoid, $u(\epsilon)/R$ reduces to $1/a$, where a is the radius of the sphere, and equations 20 and 21 reduce to 12 and 13, except for the "salting-out" term.

III

We are now ready for a brief review of the interpretation of the thermodynamic interaction of dipolar ions and electrolytes by means of the equations put forward in the preceding section. For this purpose we turn to the data of Cohn and his coworkers (1, 2, 3), relating to the influence of electrolytes on the solubility of the aliphatic amino acids and the peptides of glycine. If s and s_0 are the solubilities of a dipolar ion in a given solvent in the presence and in the absence of electrolyte, the conditions of heterogeneous equilibrium require that the ratio s_0/s be equal to γ , the activity

coefficient in the presence of electrolyte, the solution being assumed ideal when the dipolar ion component is present alone. We may therefore write a limiting solubility law in the following form,

$$\log_{10}(s/s_0) = B\Gamma \quad (22)$$

where Γ is the ionic strength and B is given by one of equations 12, 17, or 20, according to the assumed structure of the dipolar ion. The data of Cohn confirm the form of the limiting law, in which the initial term is linear in the ionic strength, and the coefficient B may be obtained from the limiting slopes of his solubility curves. In order to correlate measurements in solvents of different dielectric constant, it is convenient to introduce a coefficient B_0 equal to $(D/D_w)^2 B$, related to the solubility ratio as follows,

$$B_0 = \lim_{r \rightarrow 0} \left[\frac{d(D_w/D) \log_{10}(s/s_0)}{d(D/D_w)\Gamma} \right] \quad (23)$$

From equations 12, 17, and 20 we remark that the theory predicts that B_0 should be independent of the dielectric constant of the solvent if "salting-out" forces due to image repulsion are of small magnitude. This independence is approximately confirmed by Cohn's measurements (1) of the solubilities of the amino acids in alcohol-water mixtures, from which we may conclude that the "salting-out" forces are of secondary importance, though not negligible.

We shall first discuss the solubility data for the simplest amino acid, glycine, on the basis of the spherical dipolar ion model. The dipole moment, μ , may be computed from the solubility coefficient, B_0 , by equation 13.

$$\begin{aligned} \mu &= \sqrt{183 B_0^{(0)} a} \\ B_0^{(0)} &= (D/D_w)^2 B^{(0)} \end{aligned} \quad (24)$$

From the partial molal volume of glycine at infinite dilution in water and the ionic radii of Pauling and Huggins (9), Cohn (1) estimates a to be 3.90 for glycine and lithium chloride. From the solubility of glycine in the presence of lithium chloride in alcohol-water solvents of decreasing dielectric constant, he determines $B_0^{(0)}$ to be 0.32. Substitution of these values in equation 24 yields a value, 15 Debye units, for the glycine dipole moment. This value is in agreement with the dipole moment, calculated on the basis of structural considerations, for the glycine dipolar ion in which the terminal NH_3^+ and COO^- carry residual charges $+e$ and $-e$, respectively. It is also of interest to consider the "salting-out" coefficient $B_0^{(1)}$. Its relative importance becomes greater with increasing

dielectric constant of the solvent. Although in water the high solubility of glycine prevents the use of equation 22 without a term for the mutual dipolar ion interaction, the coefficient B_0 may be obtained from electromotive force measurements of Joseph (7) and freezing point measurements of Scatchard and Prentiss (11). For glycine and sodium chloride in water at 25°C., B_0 has the value 0.24. The difference between this value and $B_0^{(0)}$, roughly 0.08, gives the "salting-out" coefficient, $B_0^{(1)}$, on the basis of the present theory. For glycine and sodium chloride Cohn calculates α to be 4.05 and ρ is 0.7. By linear interpolation in table 1, we obtain for $\alpha(\rho)$ a value 1.37. With these values and Cohn's estimate, 57 cc., for the limiting partial molal volume of glycine, corrected for solvent electrostriction, we calculate from equation 13 the value 0.08 for $B_0^{(0)}$. Thus in water the "salting-out" contribution arising from image forces amounts to about 25 per cent of the "salting-in" contribution to the solubility coefficient, $B_0^{(0)}$, arising from the interaction of the ions of the electrolyte and the true charges of the dipolar ion.

We shall now discuss glycine and its peptides on the basis of the first ellipsoidal model, in which two charges of opposite sign are situated at the foci. The calculations will be somewhat more approximate than those for glycine on the spherical model, since the ion size of the electrolyte is neglected and the "salting-out" influence of image forces is not included. The distance R between the foci of an ellipsoid of eccentricity ϵ and volume v/N may be expressed in the following manner

$$R = \epsilon p (1 - \epsilon^2)^{-\frac{1}{2}} \quad (25)$$

where the length p is equal to $(6v/N\pi)^{\frac{1}{3}}$ and is to be expressed in Ångström units. With numerical values for the constants, we have

$$p = 1.47 v^{\frac{1}{3}} \quad (26)$$

where the molal volume v is in cubic centimeters. For a sphere p is the diameter. From equations 18, 25, and 26 we may write

$$\begin{aligned} B_0 &= 0.167 \epsilon p (1 - \epsilon^2)^{-\frac{1}{2}} g(\epsilon) \\ R &= \epsilon p (1 - \epsilon^2)^{-\frac{1}{2}} \\ p &= 1.47 v^{\frac{1}{3}} \end{aligned} \quad (27)$$

From the first of these equations and the experimental value of B_0 the eccentricity of the molecular ellipsoid may be obtained by linear interpolation in table 2. From the eccentricity and the molal volume, the

charge separation of the dipolar ion may then be computed with the second of the equations. Calculations for glycine, diglycine, triglycine, and β -alanine are summarized in table 4. They are based on Cohn's experimental values of B_0 and on his estimates of molal volumes corrected for solvent electrostriction (1). The distance, 2.8 Å., for glycine, corresponds to a dipole moment of 13, which differs only slightly from the value 15 obtained on the basis of the spherical model. A significant part of the difference between the two values is due to the neglect of electrolyte ion size in the ellipsoidal calculation. It therefore appears that we cannot conclude much as to the shape of the glycine dipolar ion from the influence of salts on its activity. However, owing to this very insensitivity to shape, we can place considerable confidence in the value of the dipole moment computed from the salt effect. The distances 4.7 and 6.4 computed for diglycine and triglycine are considerably below the values, 6.7 and 10.2, estimated by Cohn (1) for an extended chain configuration. Thus our calculations suggest that the extended chain configuration is

TABLE 4
Values of B_0 , v , and R for four dipolar ions

| | B_0 | v | R |
|-----------------------|-------|-----|-----|
| Glycine..... | 0.32 | 57 | 2.8 |
| Diglycine..... | 0.58 | 93 | 4.7 |
| Triglycine..... | 0.80 | 130 | 6.4 |
| β -Alanine..... | 0.43 | 73 | 3.6 |

not the preferred one, but that, owing to internal rotation, the average separation of the charges in these dipolar ions lies intermediate between the extended chain value and the free rotation value. Under these circumstances the computed distances have only formal significance as average distances unless a single preferred configuration should happen to dominate all others in probability. To take internal rotation properly into account, we should compute a B_0 for each internal configuration and then average over all configurations with an appropriate distribution function for comparison with the experimental value of B_0 . At present it is not possible to do this; moreover, the ellipsoidal model would be an extremely rough approximation for "crumpled" configurations.

Finally we shall discuss a series of the aliphatic alpha-amino acids on the basis of the second ellipsoidal model in which a point dipole is located at one focus. We represent glycine by a sphere with a point dipole at its center, 2.8 Å. from the surface of the molecule, and the homologs of glycine by ellipsoids with point dipoles at a focus located at the same

distance, 2.8 Å., as in glycine, from the surface of the molecule measured along the major axis. The eccentricity of the ellipsoid is then determined by the relation

$$(1 - \epsilon)(1 - \epsilon^2)^{-\frac{1}{2}} = 2l/p \quad (28)$$

where p is defined by equation 26 in terms of the molecular volume, and l , the distance of a focus from the surface, is 2.8 Å. in the subsequent calculations. Equation 28 may be roughly solved by linear interpolation

TABLE 5
 l/p as a function of the eccentricity of the ellipsoidal cavity

| $2(l/p)$ | ϵ | $2(l/p)$ | ϵ |
|----------|------------|----------|------------|
| 1.00 | 0.00 | 0.46 | 0.60 |
| 0.81 | 0.20 | 0.38 | 0.70 |
| 0.70 | 0.33 | 0.28 | 0.80 |
| 0.55 | 0.50 | | |

TABLE 6
Calculated values of the dipole moment

| ACID | B_0 | p | DIPOLE MOMENT |
|----------------------------------|-------|-----|---------------|
| Glycine..... | 0.32 | 5.6 | 13 |
| α -Alanine..... | 0.31 | 6.1 | 13 |
| α -Aminobutyric acid..... | 0.31 | 6.6 | 13 |
| Leucine..... | 0.30 | 7.3 | 13 |

in table 5. From equations 21 and 25 the following relation between the dipole moment of the dipolar ion is obtained

$$\mu = \sqrt{183B_0sp(1 - \epsilon^2)^{-\frac{1}{2}}[u(\epsilon)]^{-1}} \quad (29)$$

Calculations based upon equation 29 and Cohn's values of B_0 and molecular volumes for glycine, α -alanine, α -aminobutyric acid, and leucine are listed in table 6. The discrepancy between the dipole moment values 13 and 15, both based upon the spherical model, is merely due to the neglect of electrolyte ion size in the present calculation, a procedure which we see does not lead to great error. It is interesting that the calculated dipole moments of glycine and of the other aliphatic alpha-amino acids turn out to have identical values. This is essentially what would be expected on the basis of structural considerations, although small differences could possibly arise from induction effects in the different aliphatic chains.

The problem of internal rotation does not enter explicitly into the determination of the dipole moments of the alpha-amino acids as it did in the case of the peptides of glycine. However, the average configuration of the aliphatic chain to which the glycine residue is attached will determine how closely the actual molecule conforms to the ellipsoidal shape, which is at best an approximation.

REFERENCES

- (1) COHN, E. J.: Naturwissenschaften **20**, 663 (1932); Ann. Rev. Biochem. **4**, 93 (1935); Chem. Rev. **19**, 241 (1936).
- (2) COHN, E. J., McMEEKIN, T. L., EDSELL, J. T., AND BLANCHARD, M. H.: J. Am. Chem. Soc. **56**, 784 (1934); J. Biol. Chem. **100**, Proc. XXVIII (1933).
- (3) COHN, E. J., McMEEKIN, T. L., GREENSTEIN, J. P., AND WEARE, J. H.: J. Am. Chem. Soc. **58**, 2365 (1936).
- (4) McMEEKIN, T. L., COHN, E. J., AND BLANCHARD, M. H.: J. Am. Chem. Soc. **59**, 2717 (1937).
- (5) FERRY, R. M., COHN, E. J., AND NEWMAN, E. S.: J. Am. Chem. Soc. **60**, 1480 (1938).
- (6) FUOSS, R.: J. Am. Chem. Soc. **56**, 1027 (1934); **58**, 982 (1936).
- (7) JOSEPH, N. R.: J. Biol. Chem. **111**, 479, 489 (1935).
- (8) KIRKWOOD, J. G.: J. Chem. Phys. **2**, 351 (1934); Chem. Rev. **19**, 275 (1936).
- (9) PAULING, L., AND HUGGINS, M. L.: Z. Krust. Miner. Petrog. **A87**, 205 (1934).
- (10) SCATCHARD, G., AND KIRKWOOD, J. G.: Physik. Z. **33**, 297 (1932).
- (11) SCATCHARD, G., AND PRENTISS, S. S.: J. Am. Chem. Soc. **56**, 1486, 2314 (1934)

APPENDIX

We shall present here some of the mathematical details in the calculation of V_{ik} , the electrostatic work required to bring a dipolar ion i and a real ion k from infinite separation in the pure solvent to a given relative configuration. The real ion is represented by a point charge $z_k e$ and the dipolar ion by a charge distribution $e_1 \dots e_s$ located in a cavity ω_0 , of dielectric constant D_0 , in the solvent of dielectric constant D . The potential V_{ik} is then given by

$$V_{ik} = W - W_0 \quad (30)$$

where W is the work of charging the system in the given configuration and W_0 the work of charging when the two ions are infinitely separated. W is to be calculated by means of the formula

$$W = \frac{1}{2} \left\{ z_k e \psi_e(r_k) + \sum_{i=1}^s e_i \psi_i(r_i) \right\} \quad (31)$$

where $\psi_i(r_i)$ is the electrostatic potential in the interior of ω_0 at the point r_i of location of charge e_i and $\psi_e(r_k)$ is the potential exterior to ω_0 at the

point of location r_k of the real ion. The potentials ψ_i and ψ_o both satisfy Laplace's equation,

$$\begin{aligned}\nabla^2 \psi_i &= 0 \\ \nabla^2 \psi_o &= 0\end{aligned}\tag{32}$$

as well as the boundary conditions

$$\begin{aligned}\psi_i(\mathbf{r}) &= \psi_o(\mathbf{r}) \\ D_0 \mathbf{n} \cdot \nabla \psi_i(\mathbf{r}) &= D \mathbf{n} \cdot \nabla \psi_o(\mathbf{r})\end{aligned}\tag{33}$$

on the surface of the cavity ω_0 .

Sphere

When the cavity ω_0 is a sphere of radius b , it is convenient to employ polar coördinates (r, ϑ, φ) with origin at the center of the sphere. Potentials ψ_i and ψ_o satisfying Laplace's equation and possessing the appropriate singularities are

$$\begin{aligned}\psi_i &= \sum_{l=1}^s \frac{e_l}{D_0 |\mathbf{r} - \mathbf{r}_l|} + \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} B_{nm} r^n P_n^m(\cos \vartheta) e^{im\varphi} \\ \psi_o &= \frac{z_k e}{D |\mathbf{r} - \mathbf{r}_k|} + \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \frac{A_{nm}}{r^{n+1}} P_n^m(\cos \vartheta) e^{im\varphi}\end{aligned}\tag{34}$$

where the $P_n^m(\cos \vartheta)$ are the associated Legendre functions of the first kind. On the boundary of the sphere b , we have

$$\begin{aligned}\psi_o(b, \vartheta, \varphi) &= \psi_i(b, \vartheta, \varphi) \\ D \left(\frac{\partial \psi_o}{\partial r} \right)_{r=b} &= D_0 \left(\frac{\partial \psi_i}{\partial r} \right)_{r=b}\end{aligned}\tag{35}$$

for all values of ϑ and φ in the intervals 0 to π and 0 and 2π . On the surface of the sphere we may employ the following harmonic expansions

$$\begin{aligned}D \frac{z_k e}{|\mathbf{r} - \mathbf{r}_k|} &= \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} F_{nm} r^n P_n^m(\cos \vartheta) e^{im\varphi} \\ F_{nm} &= \frac{z_k e}{Dr_k^{n+1}} \frac{(n - |m|)!}{(n + |m|)!} P_n^m(\cos \vartheta_k) e^{-im\varphi_k} \\ \sum_{l=1}^s \frac{e_l}{D_0 |\mathbf{r} - \mathbf{r}_l|} &= \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \frac{G_{nm}}{r^{n+1}} P_n^m(\cos \vartheta) e^{im\varphi} \\ G_{nm} &= \frac{1}{D_0} \frac{(n - |m|)!}{(n + |m|)!} \sum_{l=1}^s e_l r_l^n P_n^m(\cos \vartheta_l) e^{-im\varphi_l}\end{aligned}\tag{36}$$

By substitution of equations 34 and 36 in equation 35 and use of the orthogonality of the functions

$$P_n^m(\cos \vartheta) e^{im\varphi}$$

on the surface of the sphere, we obtain the following set of linear equations for the coefficients A_{nm} and B_{nm}

$$\begin{aligned} A_{nm} + b^{2n+1} F_{nm} &= G_{nm} + b^{2n+1} B_{nm} \\ (n+1)A_{nm} - nb^{2n+1} F_{nm} &= \sigma[(n+1)G_{nm} - nb^{2n+1} B_{nm}] \quad (37) \\ \sigma &= D_0/D \end{aligned}$$

Solution yields

$$\begin{aligned} B_{nm} &= \frac{2n+1}{n+1+n\sigma} F_{nm} + \frac{(n+1)(\sigma-1)}{n+1+n\sigma} \frac{G_{nm}}{b^{2n+1}} \\ A_{nm} &= \frac{n(1-\sigma)b^{2n+1}}{n+1+n\sigma} F_{nm} + \frac{(2n+1)\sigma}{n+1+n\sigma} G_{nm} \quad (38) \end{aligned}$$

Use of equations 34 and 38 in equations 30 and 31 and application of the addition theorem of spherical harmonics yield, when the dipolar ion has no net charge,

$$\begin{aligned} V_{ik} &= \frac{z_k e}{Dr_k} \sum_{n=1}^{\infty} \sum_{l=1}^{\infty} \frac{(2n+1)e_l}{n+1+n\sigma} \left(\frac{r_i}{r_k} \right)^n P_n(\cos \vartheta_{kl}) \\ &\quad + (1-\sigma) \frac{z_k^2 e^2 b}{2Dr_k^2} \sum_{n=1}^{\infty} \frac{n}{n+1+n\sigma} \left(\frac{b}{r_k} \right)^{2n} \quad (39) \end{aligned}$$

where r_k is the distance of the real ion $z_k e$ from the center of the sphere, and ϑ_{kl} is the angle between the vectors \mathbf{r}_k and \mathbf{r}_l from the center of the sphere terminating in the real ionic charge $z_k e$ and the charge e_l of the dipolar ion. When the dipolar ion contains a point dipole at the center, the first sum in equation 39 degenerates into

$$\frac{z_k e \mu \cos \vartheta_k}{(2D + D_0)r_k^2} \quad (40)$$

and equation 11 results at once from equation 39, with a slight change in the summation index in the second sum. In this case the dipolar ion may be regarded as possessing two charges $+e$ and $-e$ situated at equal distances r_0 from the center, with $\vartheta_{k2} = \pi - \vartheta_{k1}$. If we pass to the limit $r_0 = 0$ with

$$\mu = 2 \lim_{r_0 \rightarrow 0} (er_0)$$

all terms except for $n = 1$ vanish in first sum of equation 39, the first term reducing to equation 40 if ϑ_{k1} is designated simply by ϑ_k .

Before leaving the spherical case, it is perhaps desirable to give the general expression for $B_i^{(0)}$ when the charge distribution of the dipolar ion is arbitrary. $B_i^{(1)}$ is still given by equation 13.

$$B_i^{(0)} = \frac{4\pi Ne^2}{2303(DkT)^2} \sum_{n=1}^{\infty} \frac{(2n+1)}{(2n-1)(n+1+n\sigma)} \frac{M_n}{a^{2n-1}}, \quad (41)$$

$$M_n = \sum_{l,l'} e_l e_{l'} r_l^n r_{l'}^n P_n(\cos \vartheta_{l'l})$$

where $\vartheta_{l'l}$ is the angle between the vectors of length $r_{l'}$ and r_l , joining the charges $e_{l'}$ and e_l to the center of the sphere. For a molecule of the type of cystine, a model, consisting of two point dipoles of moment μ perpendicular to a common diameter and each situated at a distance l from the center, is useful. In this case equation 41 reduces to

$$B_i^{(0)} = \frac{8\pi Ne^2}{2303(DkT)^2} \frac{\mu^2(1+\cos\varphi)}{a} \sum_{n=0}^{\infty} \frac{2n+3}{(2n+1)[n+2+(n+1)\sigma]} \left(\frac{l}{a}\right)^{2n} \quad (42)$$

where φ is the angle between the two dipole moments.

Ellipsoid

When the cavity ω_0 is ellipsoidal in form, we may conveniently employ confocal elliptical coördinates, λ , η , φ , where $\lambda = (r_1 + r_2)/R$, and $\eta = (r_1 - r_2)/R$, r_1 and r_2 being the distances of the point from the respective foci and R the interfocal distance. The angle φ measures the inclination of the plane of r_1 and r_2 to a chosen reference plane containing the major axis. The cavity ω_0 is then specified by a value λ_0 equal to the reciprocal of the eccentricity of its elliptical section. We suppose the charges $e_1 \dots e_s$ of the dipolar ion to lie on the major axis of the ellipsoid. We shall further neglect the image distribution induced in the cavity ω_0 by the real ionic charge. Potentials satisfying Laplace's equation and having the proper singularities are the following

$$\psi_e = \sum_{n=0}^{\infty} A_n P_n(\eta) Q_n(\lambda) \quad (43)$$

$$\psi_i = \sum_{l=1}^{\infty} \frac{e_l}{D_0 |\mathbf{r} - \mathbf{r}_l|} + \sum_{n=0}^{\infty} B_n P_n(\eta) Q_n(\lambda)$$

where the $P_n(\lambda)$ and $Q_n(\lambda)$ are the Legendre functions of the first and second kinds, respectively.

The boundary conditions are

$$\begin{aligned} \psi_e(\lambda_0, \eta) &= \psi_i(\lambda_0, \eta) \\ D \left(\frac{\partial \psi_e}{\partial \lambda} \right)_{\lambda=\lambda_0} &= D_0 \left(\frac{\partial \psi_i}{\partial \lambda} \right)_{\lambda=\lambda_0} \end{aligned} \quad \left. \right\} -1 \leq \eta \leq +1 \quad (44)$$

On the boundary of ω_0 the initial terms of the second of equations 43 may be developed in the Neumann expansion,

$$\sum_{i=1}^s \frac{e_i}{D_0 |\mathbf{r} - \mathbf{r}_i|} = \frac{2}{R} \sum_{n=0}^{\infty} (2n+1) G_n P_n(\eta) Q_n(\lambda_0) \\ G_n = \sum_{i=1}^s e_i P_n(\eta_i) \quad (45)$$

where $(1, \eta_i)$ are elliptical coördinates of the dipolar ionic charge e_i . Application of the boundary conditions 44 to equations 43 and 45, use of the orthogonality of the functions $P_n(\eta)$, yields the set of linear equations,

$$A_n Q_n(\lambda_0) = B_n P_n(\lambda_0) + (2/R)(2n+1) G_n Q_n(\lambda_0) \quad (46)$$

$$A_n Q'_n(\lambda_0) = \sigma [B_n P'_n(\lambda_0) + (2/R)(2n+1) G_n Q'_n(\lambda_0)]$$

where $Q'_n(\lambda_0)$ and $P'_n(\lambda_0)$ are the first derivative of the indicated functions, and σ is the ratio D_0/D . Solution of the equations 46 and elimination of the derivative $Q'_n(\lambda_0)$ by means of the formula

$$P_n Q'_n - P'_n Q_n = (1 - \lambda^2)^{-1}$$

yield

$$A_n = \frac{2\sigma}{R} \frac{(2n+1)G_n}{1 + (\lambda_0^2 - 1)(1 - \sigma)P'_n(\lambda_0)Q_n(\lambda_0)} \quad (47)$$

$$B_n = \frac{2(\sigma - 1)}{R} \frac{(2n+1)G_n}{1 - \sigma[P'_n(\lambda_0)Q_n(\lambda_0)]/[P_n(\lambda_0)Q'_n(\lambda_0)]}$$

We now calculate V_{ik} by means of the formula

$$V_{ik} = z_k e \psi_i(\lambda_k, \eta_k) \quad (48)$$

and obtain, neglecting σ in comparison with unity in the denominators on the right-hand side of equation 47,

$$V_{ik} = \frac{2z_k e}{DR} \sum_{n=1}^{\infty} \frac{(2n+1)G_n P_n(\eta_k)Q_n(\lambda_k)}{1 + (\lambda_0^2 - 1)P'_n(\lambda_0)Q_n(\lambda_0)} \quad (49)$$

The sum begins with n equal to unity since G_0 , the total charge of the dipolar ion, vanishes. By a simple algebraic transformation equation 49 may be written as follows

$$V_{ik} = \frac{2z_k e}{DR[1 + (\lambda_0^2 - 1)Q_1(\lambda_0)]} \sum_{n=1}^{\infty} (2n+1)G_n P_n(\eta_k)Q_n(\lambda_k) + Y(\lambda_k, \eta_k) \quad (50)$$

$$Y(\lambda_k, \eta_k) = \frac{2(\lambda_0^2 - 1)z_k e}{DR} \cdot \sum_{n=2}^{\infty} \frac{(2n+1)G_n [Q_1(\lambda_0) - P'_n(\lambda_0)Q_n(\lambda_0)]}{[1 + (\lambda_0^2 - 1)P'_n(\lambda_0)Q_n(\lambda_0)][1 + (\lambda_0^2 - 1)Q_1(\lambda_0)]} P_n(\eta_k)Q_n(\lambda_k)$$

The term $Y(\lambda_k, \eta_k)$ vanishes for the limiting eccentricities zero and unity of the ellipsoidal cavity ω_0 and can be neglected without great error for intermediate eccentricities. We therefore have, approximately,

$$V_{ik} = \frac{2z_k e}{DR[1 + (\lambda_0^2 - 1)Q_1(\lambda_0)]} \sum_{n=1}^{\infty} (2n + 1)G_n P_n(\eta_k)Q_n(\lambda_k) \quad (51)$$

When the charge distribution of the dipolar ion consists of two charges $+e$ and $-e$ situated at the foci $(1, 1)$ and $(1, -1)$, respectively, we have

$$G_n = e[1 - (-1)^n] \quad (52)$$

and the series in equation 51 may be summed to give equation 15, when we note that $1 + (\lambda_0^2 - 1)Q_1(\lambda_0)$ is equal to

$$\lambda_0 \left[1 - \frac{\lambda_0^2 - 1}{2} \log \frac{\lambda_0 + 1}{\lambda_0 - 1} \right]$$

On the other hand, when the charge distribution consists of a point dipole of moment μ at the focus $(1, -1)$, the G_n 's have the form

$$G_n = \lim_{x \rightarrow 0} e \left[P_n \left(-1 + \frac{2x}{R} \right) - (-1)^n \right] \quad (53)$$

$$\mu = \lim_{x \rightarrow 0} (ex)$$

and calculation yields

$$G_n = (-1)^n n(n + 1)\mu/R \quad (54)$$

Summation in equation 51 with the G_n of equation 54 yields equation 19.

THE APPLICATION OF THE THEORY OF ABSOLUTE REACTION RATES TO PROTEINS¹

HENRY EYRING

Department of Chemistry, Princeton University, Princeton, New Jersey

AND

ALLEN E. STEARN

Department of Chemistry, University of Missouri, Columbia, Missouri

Received January 30, 1939

INTRODUCTION

In discussing protein kinetics we shall make use of quantities related in the following equation (14)

$$k' = \kappa \frac{kT}{h} \exp(-\Delta F^\ddagger/RT) = \kappa \frac{kT}{h} \exp(-\Delta H^\ddagger/RT) \exp(\Delta S^\ddagger/R) \quad (1)$$
$$= \kappa \frac{kT}{h} K^\ddagger \frac{\gamma_n}{\gamma^\ddagger}$$

Here k' is the specific reaction velocity constant in reciprocal seconds; κ is a transmission coefficient (taken as unity in this paper); k is the Boltzmann constant; h is Planck's constant; T is the absolute temperature; ΔF^\ddagger is the free energy increase when the activated complex is formed from the reactant; ΔH^\ddagger and ΔS^\ddagger are the analogous increases in heat and entropy; K^\ddagger is the equilibrium constant between activated complex and reactant (both taken as ideal solutes); and γ_n and γ^\ddagger are the activity coefficients of reactants and activated complex, respectively. If there is more than one reactant, then γ_n is the product of the activity coefficients. The same generalized interpretation applies to γ^\ddagger .

From equation 1 we have

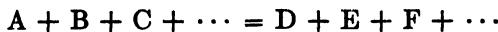
$$\Delta H^\ddagger = RT^2 \frac{d \ln k'}{dT} - RT \quad (2)$$

and

$$\Delta S^\ddagger = \frac{\Delta H^\ddagger - \Delta F^\ddagger}{T}$$

¹ Presented at the Symposium on the Physical Chemistry of Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

If the activated complex D is formed by the reaction



the concentration of D is

$$(D) = \frac{K^{\ddagger}(A)(B)(C)\gamma_A\gamma_B\gamma_C\dots}{(E)(F)\gamma_D\gamma_E\gamma_F\dots}$$

The rate of reaction is then obtained by multiplying this concentration of D by the frequency of flying apart, kT/h , and the transmission coefficient, κ , i.e., the chance that flying apart will successfully end in reaction. This gives:

$$(D)\kappa \frac{kT}{h} = K^{\ddagger} \frac{(A)(B)(C)\dots\gamma_A\gamma_B\gamma_C\dots}{(E)(F)\dots\gamma_D\gamma_E\gamma_F\dots} \kappa \frac{kT}{h}$$

Dividing out the concentrations on the right of this expression, we obtain the specific reaction rate constant as given in equation 1.

Whether the solvent should be thought of as taking part in the equilibrium leading to D depends on how the standard state is defined, just as in ordinary equilibria. Whenever activity coefficients vary rapidly with concentration, it is an indication that a strong interaction between solvent and reactants has been ignored. An alternative procedure is to redefine the standard state in such a way as to include the solvent as a reactant in the equilibrium. Thus the activity of hydrochloric acid at 1 molal concentration is 0.676, while at 16 molal it has become 478,000 (20). This of course is because of a hydration in the dilute solution which is no longer possible in the concentrated solution. It, however, illustrates the impossibility of ignoring the solvent in choosing a single standard state with reasonably constant activity coefficients in such reactions. The same difficulty is sometimes encountered in rate problems.

Part of the difficulty in the productive use of the theory of absolute reaction rates applied to proteins lies in the paucity of reactions characteristic of proteins as such. Reactions of the nature of acetylation, etc., are characteristic of fortuitous groups in particular proteins, but not of a protein. One can imagine a substance which would not undergo any specified reaction of this type but which would yet be a good protein.

Characteristic of proteins as a class are the reactions of hydrolysis and, fairly generally, denaturation. Hydrolysis of a protein is in large measure the hydrolytic splitting of the peptide bond. The theory of absolute reaction rates shows that the quantity that governs rates is the free energy of activation rather than the heat of activation alone, so that it is quite

conceivable that, even though the latter quantity be equal in several cases, the rates may be quite different because of a difference in the entropy change.

In this connection the most unequivocal data are those of Escolme and Lewis (12), who measured the rates of hydrolysis of acetylglycine and benzoylglycine. The values of ΔH^\ddagger , the heat of activation, are almost identical in the two cases (21,184 and 21,084 cal. at unit hydrogen-ion activity), but the values of ΔS^\ddagger , the entropy of activation (-24.8 and -31.8 cal. per degree, respectively), give reaction rates differing by over thirty-fold. This difference of 7 units in the entropy of activation has been accounted for by the difference in the loss of rotational entropy of internal degrees of freedom on the part of the respective substances when the activated complex is formed (14d).

In general, however, the kinetics of protein hydrolysis are fairly regular, and it is in the case of protein denaturation that the theory of absolute reaction rates exhibits its power as a tool. Denaturation of proteins is characterized by quite abnormal values of activation heats and yet is regularly of the first order with respect to protein.

Recently (3, 26, 14c) the process of denaturation has been pictured as the simultaneous breaking of a number of weak bonds in the molecule with a resulting loss of structure, or, what amounts to the same thing, a large increase in randomness. The fact of finite rates coupled with large heats of activation necessarily implies a large increase of entropy with activation.

Steinhardt (35) and La Mer (18), on the basis of the work of Steinhardt on the inactivation of pepsin, have considered the matter from a somewhat different point of view. Steinhardt's data show that, through a restricted pH range, the rate of pepsin inactivation near the neutral point increases almost exactly as the fifth power of the hydroxide-ion activity. Enzyme inactivation proves that some particular group or groups have disappeared but leaves us in the dark as to what further changes accompany it. This pH dependence indicates, as these authors point out, that the reactive form in this reaction is an ion quintuply more ionized than the stable form. However, to get what they term the true heat of activation, the usually accepted value, i.e.,

$$RT^2 \frac{d \ln k'}{dT} - RT$$

is corrected by subtracting the heat necessary to bring about this quintuple ionization.

An analysis of Steinhardt's results is given in table 1. Here the symbol P_0 refers to the protein before the five protons have been lost to transform it into P_5 . By a further activation of P_5 the activated complex P_5^\ddagger is

formed. The values in the first row are calculated directly from the measured reaction rates at a pH where most of the protein is in the form P_0 . We assume, following Steinhardt, five equal dissociation constants, K_0 , whose values are 1.74×10^{-7} at 25°C . and 1.024×10^{-7} at 15°C . In row 3, K' is the value of $(P_s)/(P_0)$ at pH 5.7 and is equal to $(K_0/a_{\text{H}})^5$. From the values in rows 1 and 3 those in row 2 follow at once.

It is apparent that at pH 5.7 almost all of the activation entropy of P_0 going to P_s^t is associated with the preliminary ionization of protons into solution to give P_s . This large entropy increase is due to the fact that the protons are ionizing into a solution of very low hydrogen-ion activity. When ionization takes place into a solution of unit hydrogen-ion activity, we note that there is little or no entropy of ionization (row 4). This last way of calculating a specific rate constant by including explicit account of

TABLE 1
The kinetics of pepsin inactivation

| pH | PROCESS | LOG ₁₀ k' sec. ⁻¹ | | ΔH^\ddagger calories | ΔF^\ddagger calories | ΔS^\ddagger |
|-----|--------------------------------------------------------------------------------------|-----------------------------------------|--------|---------------------------------|---------------------------------|---------------------|
| | | 25°C. | 15°C. | | | |
| 5.7 | $P_0 \rightarrow P_s^t$ | -6.996 | -8.716 | 67,470 | 26,955 | 135.9 |
| | $P_s \rightarrow P_s^t$ | -1.696 | -2.266 | 22,360 | 19,737 | 8.8 |
| | | LOG ₁₀ K' | | ΔH | ΔF | ΔS |
| 5.7 | $P_0 \rightarrow P_s$ | -5.3 | -6.45 | 45,110 | 7,218 | 127.1 |
| 0 | $P_0 + 5\text{H}_2\text{O} \rightarrow P_s + 5\text{H}_3\text{O}$ | -33.8 | -34.95 | 45,110 | 46,030 | -3.1 |
| 0 | $5\text{HOAc} + 5\text{H}_2\text{O} \rightarrow 5\text{OAc}^- + 5\text{H}_3\text{O}$ | -23.75 | | -485 | 32,350 | -110.2 |

the hydrogen-ion activity is the accepted procedure of adopting a standard state used for ordinary equilibria.

It is highly desirable, wherever possible, to adopt the same standard states for defining reaction rate constants as for equilibria. Reaction rate constants if properly measured, i.e., with due regard to activity coefficients of reactants and activated complex, are functions only of temperature and pressure. However, with the data at present available this is usually impossible, so that the calculated values of ΔF^\ddagger , ΔH^\ddagger , and ΔS^\ddagger given refer, in general, to the unspecified standard state under which the measurements happen to be made. Sometimes certain reactants are even omitted in calculating the specific reaction rate constant, k' . For such a k' the standard state of the omitted reactant is the activity at which it happened to be present in the solution. Some, although by no means all, of the abnormally large activation entropies of protein reactions are due to failure

to include the hydrogen-ion concentration properly in calculating the specific reaction rate constants.

Row 2 in table 1 shows an apparent entropy of activation of P_s of only 8.8 units, so that the amount of "opening up" of the protein molecule seems at first glance insignificant. It is, however, apparent that the protein in the preceding ionization acts as a very peculiar acid. When we compare the entropy involved in ionizing five moles of acetic acid, we find -110 units in place of -3. This difference indicates that the actual increase in the entropy of "opening up" the protein molecule may be in the neighborhood of 100 units, when P_t^{\ddagger} is formed from P_s . The opening up of the protein structure measured by this entropy increase is probably associated with the breaking of salt bridges between groups such as $-\text{NH}_3^+$ and $-\text{COO}^-$.

Mirsky and Pauling (26) assumed that such acid and basic groups were held together by hydrogen bonds which were broken by denaturation. Sookne and Harris (34) find that salt bonds do not contribute greatly to the strength of wool. This is at least consistent with our view that in the process studied by Steinhardt the chief contribution to the free energy of activation is in the breaking of covalent linkages less influenced by pH than the above electrovalent salt bridges.

Cohn, McMeekin, Edsall, and Blanchard (10) have considered the possibility of salt-bond formation in proteins. They were, however, concerned with the additional bonds which can form in the denatured protein as a result of the greater flexibility consequent upon loss of structure. The salt bonds here considered are those which stabilize the native protein with respect to the activated complex. Dr. Anson² points out that stabilization by salt bonds involving free amino groups cannot be of primary importance, since substances such as formaldehyde, although they react with the amino group reducing the basicity, do not greatly reduce the stability of the native protein.

The apparent entropy of activation for denaturation passes through a maximum at a pH of maximum stability. On the alkaline side the decrease in ΔS^{\ddagger} below the maximum value is due to a change in process. This is brought out in table 1 where, as the pH increases, the activation process will be seen to change from $P_o \rightarrow P_t^{\ddagger}$ with $\Delta S^{\ddagger} = 135.9$ units to the process $P_s \rightarrow P_t^{\ddagger}$, with $\Delta S^{\ddagger} = 8.8$ units. On the other side of the pH of maximum stability, as the pH decreases, the same process can take place with values of ΔS^{\ddagger} decreasing from 127.1 at pH 5.7 to -3.1 at pH 0.

The final step in activation lies in the breaking of one or more bridges which are not destroyed by ionization, such as a covalent bonded cystine bridge. Such a process cannot "pay its way" with an increase in entropy

² Private communication.

compensating for the higher energy, since it probably involves hydrolysis of the S—S bond, at least as a first step.

Salt bridges and covalent cross linkages particularly due to cystine have been frequently spoken of (37, 5, 8, 10).

The behavior of cystine itself fits this picture well. Lavine (19) finds an easy hydrolysis of cystine into cysteine and sulfinic acid when catalyzed by mercuric ion. The ease of reversibility of this dismutation is suggestive in connection with the reversibility of denaturation. Lavine found a 90 to 92 per cent cystine recovery on removing the mercury. Shinozaki and Kilpatrick (33) noted a spontaneous dismutation of cystine to cysteine and oxy acids under proper conditions. From their semi-quantitative data we obtain a value for ΔH^\ddagger of the order of 20,000 cal. By analogy with hydrolysis of peptide linkages (12) we expect about this value for ΔH^\ddagger and also a negative ΔS^\ddagger of the order of -20 to -30 units. The fact that in the case of pepsin there is an actual increase in entropy of 8 e.u. may well arise from the consequent loosening of the protein structure.

It is interesting to inquire why the activated complex should be more highly ionized than the normal pepsin. The explanation follows immediately from the concept that the activated complex contains broken salt bridges and as a result approximates the isoelectric state over a much shorter pH range. The reason is that, since the positive and negative groups which in the normal state formed the salt bridges are no longer adjacent, it becomes easier in alkaline solution for the amino or for the imino group to lose its proton and likewise, in the more acid range, easier for the negative carboxyl group to acquire a proton.

If the breaking of a salt bridge is, as we believe, an important step in the activation process leading to denaturation, there is a critical pH on either side of which salt bridges are automatically broken. The pH of maximum stability will not in general coincide with the isoelectric point. Thus if a protein contains a large excess of acid groups and a low isoelectric point, the pH of maximum stability will be higher than the isoelectric point. This is well exemplified in the cases of pepsin (isoelectric point, 2.85; pH of maximum stability, 5.28) and of egg albumin (isoelectric point, 4.55; pH of maximum stability, 6.76) (28). By the same token, if basic groups are largely predominant, the relation between these two points should be reversed. With a protein like hemoglobin, the isoelectric point of which is very close to neutrality and which probably has nearly the same number of acid and basic bridge-forming groups, the pH of maximum stability should practically coincide with the isoelectric point. Table 3 shows that this is the case. Proteins with very acid isoelectric regions will, according to this rule, be expected to have maximum stabilities at more alkaline pH values, since their comparatively acid isoelectric points arise from an excess of acid over basic groups.

An instructive example of the fact that the important quantity in activation is the free energy rather than the heat is given by the case of trypsin. From Pace's data (29) we get a value for ΔH^\ddagger of activation of 40,000 cal. at pH 6.5, while Anson and Mirsky (2), from thermal equilibrium data at pH 2 to 3, find ΔH for the equilibrium to be 67,600 cal. Unless there is a change of nearly 30,000 cal. in these values in the pH range between the two sets of measurements, one is led to a negative value for the heat of activation for recovery of enzymatic activity. At 50°C., however, this latter rate is seven times smaller than the rate for the loss of enzymatic activity. Any cutting down of Pace's value for the heat of activation because of the difference in pH for the two experiments would lead to a still more negative heat of activation for the recovery of enzymatic activity. Such a reaction, interpreted in terms of collisions between water and enzyme molecules, would involve an exceedingly small steric factor.

Anson and Mirsky find a total entropy increase of 213.1 units for the equilibrium, whereas Pace's data on activation show that the activated complex has only 44.7 entropy units more than active trypsin. This means that the formation of the activated complex from the inactive trypsin, for the reaction in the reverse direction, must take place with the loss of $213.1 - 44.7$ or 168.4 E.U. Thus the free energy of this reverse reaction is not negative but has at 50°C. the value 26,800 cal. ($\Delta H^\ddagger - T\Delta S^\ddagger = 40,000 - 67,600 + 323 \times 168.4$), giving a normal reaction rate. It will be seen below (table 3) that a negative temperature coefficient is obtained for the denaturation of egg albumin by urea. In such a case the ordinary collision viewpoint is less illuminating than thinking in terms of an equilibrium between the activated complex and normal molecules. In the latter case we draw on whatever knowledge there is of analogous equilibria. More accurate trypsin rate measurements will probably alter somewhat the above provisional calculations.

It is unfortunate that there are no other data analogous to those for trypsin discussed above. The results of Anson and Mirsky (1) on chemical denaturation involve the heat of combination of the protein with the denaturing catalyst, the value of which is not known. The unusual fact that ΔH for equilibrium exceeds ΔH^\ddagger for activation means, of course, that after activation there are many bonds that open spontaneously, because for them the net value of $T\Delta S$ exceeds ΔH .

II

A general view of the kinetics of protein denaturation (and enzyme or hormone inactivation) is presented in tables 2 and 3. Table 2 includes for the most part data for ordinary water denaturation (or inactivation).

These data bring out in a very striking manner the fact that the apparent heat of activation, obtained from the temperature coefficient, while important, is not in itself the dominating factor. Specifically one may point out two reactions at 50°C. (those of pancreatic proteinase and goat hemolysin), with heats varying between 38,000 and 198,000 cal. but with rates which are equal to within a factor of less than 2. The ratio of these rates, if the entropy were the same in the two cases, should be 10^{107} at 50°C.

TABLE 2
Kinetics of denaturation and enzyme inactivation

| SUBSTANCE | <i>t</i> | <i>k'</i> SEC. ⁻¹ | ΔH^\ddagger | ΔS^\ddagger | ΔF^\ddagger | $T\Delta S^\ddagger$ | pH |
|------------------------------|----------|------------------------------|---------------------|---------------------|---------------------|----------------------|------|
| | °C. | | calories | calories | calories | | |
| Insulin (16)* | 80. | $1. \times 10^{-4}$ | 35,600 | 23.8 | 27,200 | 8,400 | 1.5 |
| Pepsin (36) | 25. | 5.47×10^{-4} | 55,600 | 113.3 | 21,900 | 33,750 | 6.44 |
| Leucosin (23) | 55. | 1.2×10^{-3} | 84,300 | 185. | 23,620 | 60,680 | 6.1 |
| Egg albumin (21) | 65. | 2.54×10^{-4} | 132,100 | 315.7 | 25,400 | 106,700 | 5.0 |
| Hemoglobin (21) | 60.5 | 4.3×10^{-4} | 75,600 | 152.7 | 24,700 | 50,900 | 5.7 |
| Trypsin (29) | 50. | 2.83×10^{-5} | 40,160 | 44.7 | 25,700 | 14,440 | 6.5 |
| Enterokinase (33) | 50. | 7.37×10^{-5} | 42,160 | 52.8 | 25,100 | 17,050 | |
| Trypsin kinase (34) | 50. | 3.17×10^{-5} | 44,260 | 57.6 | 25,700 | 18,800 | |
| Proteinase (pancreatic) (35) | 50. | 1.27×10^{-4} | 37,860 | 40.6 | 24,750 | 13,100 | |
| Lipase (pancreatic) (25) | 50. | 1.18×10^{-3} | 45,360 | 68.2 | 23,330 | 22,030 | |
| Hemolysin (goat) (15) | 50. | 2.09×10^{-4} | 198,000 | 537. | 24,550 | 173,500 | |
| Vibriolysin (15) | 50. | $3. \times 10^{-3}$ | 127,950 | 326. | 22,650 | 105,300 | |
| Tetanolysin (15) | 50. | 2.1×10^{-4} | 172,650 | 459. | 24,350 | 148,400 | |
| Peroxidase (milk) (42) | 70.1 | 5.1×10^{-4} | 185,300 | 466.4 | 25,300 | 160,000 | |
| Rennin (4) | 50. | 6.7×10^{-3} | 89,350 | 208.1 | 22,130 | 67,200 | |
| Amylase (malt) (24) | 60. | $8. \times 10^{-4}$ | 41,630 | 52.3 | 24,230 | 17,400 | |
| Invertase (yeast) (13) | 50. | 7.7×10^{-5} | 52,350 | 84.7 | 25,000 | 27,350 | 5.7 |
| Invertase (yeast) (13) | 55. | $5. \times 10^{-5}$ | 86,350 | 185. | 25,700 | 60,680 | 5.2 |
| Invertase (yeast) (13) | 60.2 | 3.5×10^{-5} | 110,350 | 262.5 | 25,500 | 84,840 | 4.0 |
| Invertase (yeast) (13) | 55. | 3.75×10^{-4} | 74,350 | 152.4 | 24,400 | 50,000 | 3.0 |
| Emulsin (wet) (39) | 60. | $4. \times 10^{-3}$ | 44,930 | 65.3 | 23,200 | 21,750 | |
| Emulsin (dry) (39) | 100. | 1.48×10^{-4} | 25,550 | -7.86 | 28,500 | -2,930 | |
| Lipase (dry) (27) | 120. | $4. \times 10^{-4}$ | (24,200) | (-13.) | (29,300) | (-5,100) | |

* References to the literature are given in parentheses.

Naturally $\Delta F^\ddagger/T$ does not vary greatly for these reactions. For a conveniently measurable reaction rate between, say, 25°C. and 70°C. the free energy cannot vary much. (For example, for specific rates between 10^{-3} and 10^{-5} at temperatures between 25°C. and 70°C., ΔF^\ddagger can vary through a range of 6500 cal.) In table 2, except for the two cases of dry heating, the values of ΔF^\ddagger vary from 21,900 to 27,200, or through a range

of 5300 cal. These ΔF^\ddagger values are to be compared with the corresponding values of $T\Delta S^\ddagger$, where an entirely different situation obtains. $T\Delta S^\ddagger$ values vary over a range of more than 150,000 cal.

In many studies the pH has not been reported. It will be considered in discussing table 3, but in table 2 the data for invertase are striking. Here any variation of ΔF^\ddagger is small enough so that the slightly different temperatures at which the measurements were made obscure any noticeable trend, whereas $T\Delta S^\ddagger$ varies by over 57,000 cal. in the pH range 3 to 5.5, and passes through a distinct maximum at about pH 4.

The apparent abnormalities met with in reactions of the type summarized in table 2 are, at least in some cases, due to the method of calculating k' , a method which does not take into full account all substances involved in the equilibrium between the normal and the activated states of the reacting system. As we have previously stated, in effect it amounts to choosing a variable standard state for this equilibrium. To obtain a value of k' which is really constant (i.e., depending only on temperature and pressure), it would be necessary to include in this constant the activities of all substances which enter the equilibrium between the normal and the activated states. Suppressing any of these amounts to treating as unit activity the actual activity of the suppressed substance at which k' is measured, and thus may involve a different standard state for each measurement. It is analogous to ignoring the activities of certain of the substances entering into any equilibrium. The value of K would change with changes in the activities of these ignored substances. For example, if the constant for the ionization of phosphoric acid were formulated as $(\text{PO}_4^{3-})/(\text{H}_2\text{PO}_4^-)$, the value of this ratio would certainly change enormously when measured at widely differing hydrogen-ion activities.

However, even if we define our standard state in solution by taking the activity equal to the molality at infinite dilution, as Lewis and Randall do, there will still be reactions with abnormal entropies of activation just as one finds for the hydrogen chloride liquid-vapor equilibrium. If an activation process involves an increase or decrease in the number of ions present, the consequent "freezing" or "melting" of water around the ion will insure abnormal changes in entropy which render hopeless any attempt to treat the process by simple kinetic theory. There is, of course, no such limitation on the general statistical-mechanical treatment.

The proteins whose behavior on denaturation has been most widely studied are egg albumin and hemoglobin. In table 3 data on the kinetics of denaturation of these proteins under varying catalytic conditions are given. Specific numerical data are available, in one or the other case, for denaturation by urea, acid, alcohol, and water; data are available on denaturation by water both in the presence and in the absence of high salt

TABLE 3
Kinetics of denaturation with various catalysts

| PROTEIN AND REFERENCE | t | pH | CATALYST | K' sec. ⁻¹ | ΔH^\ddagger | ΔS^\ddagger | ΔF^\ddagger | TAS [†] |
|-----------------------|------|------|--------------------------------------------|------------------------------------------------|---------------------|---------------------|---------------------|------------------|
| Egg albumin (21)..... | 65. | 5.0 | Water | 2.54×10^{-4} | 132,100 | 315.7 | 25,400 | 106,700 |
| | 65. | 6.76 | Water | 2.8×10^{-4} | 128,200 | 295.2 | 28,400 | 99,800 |
| | 65. | 7.7 | Water | 1.8×10^{-4} | 134,300 | 317.1 | 27,100 | 107,200 |
| | 70.2 | 5.0 | Water | 5.1×10^{-3} | 132,100 | 315.7 | 23,750 | 108,350 |
| Egg albumin (21)..... | 70.2 | 6.76 | Water | 5.15×10^{-3} | 128,200 | 295.2 | 26,900 | 101,300 |
| | 70.2 | 7.7 | Water | 3.8×10^{-4} | 134,300 | 317.1 | 25,500 | 108,800 |
| | 90.5 | 5.7 | Water | 4.3×10^{-4} | 75,600 | 152.7 | 24,700 | 50,900 |
| Hemoglobin (21)..... | 60.5 | 6.8 | Water | 1.05×10^{-4} | 76,300 | 152.0 | 25,600 | 50,700 |
| | 60.5 | 8.0 | Water | 3.35×10^{-4} | 77,300 | 157.3 | 24,840 | 52,460 |
| | 68. | 5.7 | Water | 5.4×10^{-4} | 75,600 | 152.7 | 23,530 | 52,070 |
| Egg albumin (9)..... | 68. | 6.8 | Water | 1.35×10^{-3} | 76,300 | 152.0 | 24,470 | 51,830 |
| | 68. | 8.0 | Water | 4.45×10^{-3} | 77,300 | 157.3 | 23,660 | 53,640 |
| | 51.1 | 3.4 | Water | 2.76×10^{-4} | 96,750 | 223.7 | 24,300 | 72,500 |
| Egg albumin (9)..... | 56.2 | 3.4 | Water | 2.88×10^{-4} | 96,750 | 223.7 | 23,100 | 73,640 |
| | 64.8 | 9.8 | Water | 1.15×10^{-4} | | | 27,500 | |
| | 70.9 | 75.4 | $2 N (NH_4)_2SO_4$, $2 N (NH_4)_2SO_4$ | 1.94×10^{-4} 1.03×10^{-4} | 87,500 | 174.2 | 27,800 | 59,900 |
| Egg albumin (9)..... | 64. | 6.76 | $0.15 N (NH_4)_2SO_4$ | 3.41×10^{-4} | 76,500 | 152.5 | 25,100 | 51,400 |
| | 70. | 6.76 | $0.15 N (NH_4)_2SO_4$ | 2.57×10^{-3} | 76,500 | 152.5 | 24,200 | 52,300 |
| | 64. | 6.76 | $1.14 N (NH_4)_2SO_4$ | 5.45×10^{-4} | 87,000 | 184.6 | 24,800 | 62,200 |
| Hemoglobin (22)..... | 70. | 6.76 | $1.14 N (NH_4)_2SO_4$ | 5.4×10^{-3} | 87,000 | 184.6 | 23,700 | 63,300 |
| | 64. | 6.76 | $2.27 N (NH_4)_2SO_4$ | 9.1×10^{-4} | 103,800 | 230.9 | 26,000 | 77,800 |
| | 70. | 6.76 | $2.27 N (NH_4)_2SO_4$ | 1.4×10^{-4} | 103,800 | 230.9 | 24,600 | 79,200 |
| Hemoglobin (22)..... | 64. | 6.76 | $3.03 N (NH_4)_2SO_4$ | 2.44×10^{-4} | 119,800 | 275.8 | 26,900 | 92,900 |
| | 70. | 6.76 | $3.03 N (NH_4)_2SO_4$ | 5.7×10^{-4} | 119,800 | 275.8 | 25,200 | 94,600 |

| | | | | | | | | |
|---------------------------------|------|------|-----------------------|------------------------|-----------|---------|----------|-----------|
| Egg albumin (17) . . . | 0. | 5.9 | Urea, 0.6 g. per cc. | 1.7×10^{-4} | -6,750 | -95.6 | 19,350 | -26,100 |
| | 23. | 5.9 | Urea, 0.6 g. per cc. | $7. \times 10^{-4}$ | -6,780 | -95.8 | 21,560 | -28,350 |
| Hemoglobin (7, 21) . . . | 60.5 | 6.0 | Alcohol (none) | 2.3×10^{-4} | 76,250 | 153.4 | 25,100 | 51,150 |
| | 45. | 6.0 | Alcohol (20 volume %) | 4.56×10^{-4} | 107,400 | 263.9 | 23,500 | 83,900 |
| | 33. | 6.0 | Alcohol (30 volume %) | 5.3×10^{-4} | 117,000 | 308.9 | 22,500 | 94,500 |
| | 60.5 | 6.5 | Alcohol (none) | 1.22×10^{-4} | 75,050 | 148.5 | 25,550 | 49,500 |
| Hemoglobin (7, 21) . . . | 45. | 6.5 | Alcohol (20 volume %) | 4.5×10^{-4} | 102,900 | 248.8 | 23,500 | 79,100 |
| | 33. | 6.5 | Alcohol (30 volume %) | 6.8×10^{-4} | 120,100 | 319.5 | 22,300 | 97,800 |
| | 60.5 | 7.0 | Alcohol (none) | 1.22×10^{-4} | 76,050 | 151.5 | 25,550 | 50,500 |
| Hemoglobin (7, 21) . . . | 45. | 7.0 | Alcohol (20 volume %) | 8.0×10^{-4} | 98,200 | 236.1 | 23,100 | 75,100 |
| | 33. | 7.0 | Alcohol (30 volume %) | (1.1×10^{-4}) | (128,500) | (348.) | (22,000) | (106,500) |
| | 45. | 4.85 | Acid | 1.7×10^{-4} | 17,760 | -19.9 | 24,100 | -6,330 |
| Hemoglobin (16) . . . | 45. | 4.23 | Acid | 2.53×10^{-4} | 10,660 | -36.9 | 22,400 | -11,730 |
| | 37. | 4.64 | Acid | 2.78×10^{-4} | 11,300 | -38.3 | 23,200 | -11,870 |
| Hemoglobin (16) . . . | 37. | 4.11 | Acid | 3.06×10^{-4} | 10,700 | -35.5 | 21,700 | -11,000 |
| | 25. | 4.52 | Acid | 2.29×10^{-4} | 11,100 | -37.8 | 22,400 | -11,260 |
| Egg albumin (16) . . . | 25. | 4.08 | Acid | 1.76×10^{-4} | 11,100 | -33.7 | 21,150 | -10,040 |
| | 37. | 2.00 | Acid | 2.15×10^{-5} | 48,400 | 76.3 | 24,750 | 23,650 |
| Egg albumin (16) . . . | 37. | 1.62 | Acid | 5.68×10^{-5} | 36,900 | 41.2 | 24,100 | 12,770 |
| | 37. | 1.32 | Acid | 8.83×10^{-5} | 36,700 | 41.4 | 23,900 | 12,830 |
| Egg albumin (16) . . . | 25. | 1.72 | Acid | 4.8×10^{-6} | 34,400 | 32.6 | 24,700 | 9,700 |
| | 25. | 1.52 | Acid | 6.4×10^{-6} | 35,400 | 36.5 | 24,500 | 10,900 |
| | 25. | 1.35 | Acid | 8.2×10^{-6} | 35,200 | 36.3 | 24,400 | 10,800 |
| | 25. | 1.02 | Acid | 1.59×10^{-6} | (35,400) | (38.5) | 24,000 | (11,400) |

TABLE 3—Concluded

| PROTEIN AND REAGENT | t °C. | pH | CATALYST | K' sec. ⁻¹ | ΔH^\ddagger | ΔS^\ddagger | ΔF^\ddagger | TAS [†] |
|-----------------------|----------|------|-----------------------|-----------------------|---------------------|---------------------|---------------------|------------------|
| Hemoglobin (16)..... | 45. | 5.5 | | 9.6×10^{-4} | | | | 25,920 |
| | 45. | 4.9 | | 1.21×10^{-4} | | | | 24,320 |
| Egg albumin (16)..... | 45. | 4.52 | | 7.26×10^{-4} | | | | 23,190 |
| | 45. | 4.23 | | 2.53×10^{-3} | | | | 22,400 |
| Egg albumin (16)..... | 45. | 3.52 | | 1.2×10^{-4} | | | | 25,780 |
| | 45. | 2.68 | | 3.5×10^{-4} | | | | 25,100 |
| Hemoglobin (7)..... | 45. | 1.98 | | 1.73×10^{-4} | | | | 24,100 |
| | 45. | 1.34 | | 3.88×10^{-4} | | | | 23,500 |
| Hemoglobin (7)..... | 45. | 6.57 | Alcohol (20 volume %) | 4.52×10^{-4} | 100,450 | 242.1 | 23,500 | 77,000 |
| | 40. | 6.57 | Alcohol (20 volume %) | 3.5×10^{-4} | 100,450 | 242.1 | 24,700 | 75,800 |
| Egg albumin (11)..... | 60.5 | 6.0 | Alcohol (2 volume %) | 5.06×10^{-4} | 83,500 | 176.7 | 24,600 | 58,950 |
| | 55. | 6.0 | Alcohol (2 volume %) | 6.0×10^{-4} | 83,500 | 176.7 | 25,600 | 57,950 |
| Egg albumin (11)..... | 25. | 1.35 | Acid | 8.2×10^{-4} | 35,100 | 36.1 | 24,350 | 10,750 |
| | 37. | 1.35 | Acid | 8.5×10^{-4} | 35,100 | 36.1 | 23,900 | 11,200 |
| | 45. | 1.35 | Acid | 3.83×10^{-4} | 36,200 | 39.7 | 23,600 | 12,600 |

concentration. The process which takes place at pH values fairly close to the pH of maximum stability we call denaturation by water. Qualitatively the experimental behavior may be summarized in the following statements:

(1) For a given reaction ΔF^\ddagger at constant pH decreases slowly but steadily with temperature except for urea denaturation.

(2) On either side of a pH of maximum stability ΔF^\ddagger decreases respectively with increasing or decreasing pH.

(3) Again excepting the case of urea denaturation, $T\Delta S^\ddagger$ increases with increasing temperature. This may be entirely due to the temperature factor, but there are indications that ΔS^\ddagger increases slightly with temperature.

(4) Owing to the possibility of confusing side reactions, such as partial digestion, data in alkaline solutions are scarce. In sufficiently acid solutions, however, ΔS^\ddagger decreases with decreasing pH. In the cases of egg

TABLE 4
Relative effectiveness of denaturing catalysts

| CATALYST | ORDER OF INCREASING VALUES OF | | |
|--------------------|-------------------------------|---------------------|----------------------|
| | ΔF^\ddagger | ΔH^\ddagger | $T\Delta S^\ddagger$ |
| Urea..... | 1 | 1 | 1 |
| Acid..... | 2 | 2 | 2 |
| Alcohol..... | 3 | 5 | 5 |
| Water..... | 4 | 4 | 4 |
| Salt solution..... | 5 | 6 | 6 |
| Dry heat..... | 6 | 3 | 3 |

albumin and pepsin, and ostensibly in general, ΔS^\ddagger also decreases with increasing pH in sufficiently alkaline solutions.

(5) It is instructive to arrange the various denaturing catalysts in the order of increasing values for the three important quantities ΔF^\ddagger , ΔH^\ddagger , and $T\Delta S^\ddagger$. This is done in table 4.

(6) The effect of alcohol and of ammonium sulfate should be noted. In the concentration ranges covered, the values of ΔH^\ddagger and $T\Delta S^\ddagger$ are almost parallel. At the maximum respective concentrations, however, ΔF^\ddagger is rapidly decreasing with increase in alcohol concentration, whereas it is increasing with increasing salt concentration. This signifies a more rapid increase in ΔS^\ddagger with alcohol concentration, where $T\Delta S^\ddagger$ is gaining on ΔH^\ddagger , than with salt concentration, where $T\Delta S^\ddagger$ is falling behind ΔH^\ddagger .

(7) Finally a certain consistency in behavior between egg albumin and hemoglobin appears. The value of ΔS^\ddagger for egg albumin decreases by some 250 units between pH 5 measured at 65°C. and pH 1.5 measured at 37°C.,

while ΔS^\ddagger for hemoglobin decreases by about 190 units between pH 5.7 measured at 60°C. and pH 4.6 measured at 37°C. The change in ΔS^\ddagger for the two substances would no doubt agree even more closely for more nearly the same change in pH.

III

From equation 1 we see that any change in environment which renders the activated complex more soluble (comparatively) than native protein, i.e., which preferentially decreases γ^\ddagger , will increase the rate of denaturation. It is at once apparent: (1) that the temperature coefficient of solubility of the activated complex in water and most aqueous solutions is greater than that of native protein (concentrated urea solution is an exception); (2) that a shift in pH in either direction from a critical value characteristic of the particular protein increases the solubility of the activated complex preferentially; (3) that its solubility is also increased preferentially by alcohol-water mixtures, at least up to 30 per cent alcohol, and by other chemical agents such as urea, salicylate, etc.; (4) that, with increasing salt concentration, the activated complex is first preferentially "salted in" and finally preferentially "salted out".

In each of the above-mentioned environments the general mechanism of activation is the same. The protein chain is held in a specific configuration by side bridges, among which salt bridges and the cystine type of bridge are essential. In this latter type we include any bridge not automatically broken by an equilibrium shift due to change in pH and not included in so-called hydrogen bridges.

Salt bridges may be broken either by neutralization due to a pH shift, in which case the heat involved will be written off as heat of neutralization and will not appear in the measured ΔH^\ddagger , or they may be broken "thermally", in which case the heat involved will be the difference between the strength of the bridge bond and the strength of any bonds subsequently formed with solvent. This net heat may at times be negative, as in the case of concentrated urea solutions, or it may reach values higher than for water, as in alcohol or in concentrated salt solutions where the water is already tied up by salt ions. So long as we treat denaturation as a monomolecular reaction, it is perfectly legitimate to speak of the ΔH^\ddagger given by equation 2 as the heat of activation for the particular experimental conditions under which it is measured.

We speak of the process of breaking a bridge, or a set of bridges, as paying its way if the total $T\Delta S^\ddagger$ for the process equals or exceeds ΔH^\ddagger , giving $\Delta F^\ddagger \gtrless 0$. In general we do not expect the breaking of a bridge of the cystine type to pay its way unless the process is accompanied by a large opening of the protein structure, since hydrolysis is probably involved,

where ΔS^\ddagger will be negative unless the remainder of the protein molecule can more than make it up.

Certain of the salt bridges cannot pay their way in breaking; others can. In general the last few bridges broken more than pay their way, since their breaking will permit complete opening up of the protein structure. The entropy change for breaking a particular bridge will be the sum of three contributing terms: (1) that involved in freeing the side chains which form the bridge, (2) that involved in freezing solvent molecules to the free side chains, and (3) that due to the effect of the breaking of the bridge on the opening up of the protein molecule. It will be realized that two samples of the bridge in a protein molecule may make entirely different contributions to the net ΔS^\ddagger . If one is among the first broken, its contribution may even be negative, whereas if the other is among the last broken, the contribution may be very large, owing to the third term listed above.

If, when breaking, all bridges paid their way under all conditions, denaturation would proceed spontaneously and rapidly to completion. Indeed, native protein would be unknown. The process of activation, then, consists of the breaking of those bridges which will not pay their way, the number of such varying under different conditions. As a specific example, egg albumin has many salt bridges at pH 6.76 and a highly restricted configuration. When it activates ΔH^\ddagger is over 130,000 cal. and ΔS^\ddagger about 300 units. When the pH is shifted to 1, all the salt bridges are broken by ionization. There are, however, two cystine residues found per molecule of egg albumin (8). It was shown above (33) that we expect a ΔH^\ddagger of about 20,000 cal. and a ΔS^\ddagger of -20 to -30 units for the hydrolysis of the S-S bond. For two cystine bridges these values agree satisfactorily with the measured value of ΔH^\ddagger for denaturing egg albumin at pH 1, i.e., about 35,000 cal. With ΔS^\ddagger we do not, of course, expect such agreement, since the opening of the protein molecule brings the measured ΔS^\ddagger of denaturation up to about 36 units. With hemoglobin the data are equally interesting, though the case is more complicated. Hemoglobin has one cystine residue per molecule of weight 68,000 (8). Its pH stability range, however, unlike that of egg albumin, is only from about 6 to 8 (38). In urea solution it goes to a molecular weight of 34,000 (8). At pH 4.1, therefore, we assume hemoglobin with molecular weight 34,000 and only a few salt bridges. (Unless the two dissociated varieties differ, there would be only the equivalent of one-half of a cystine residue per molecule of weight 34,000.) Under these conditions the heat of activation is so small (about 11,000 cal.) that the molecule can activate by breaking a very few bridges which do not pay their way (ΔS^\ddagger about -35 units). Indeed, if the pH were lowered much farther, ΔF^\ddagger would decrease so that the rate of denaturation would be immeasurably rapid. In table 3 it is seen that in the

pH range 5.5 to 4.2 ΔF^\ddagger for hemoglobin is dropping by almost 3000 cal. per pH unit, whereas, even in the lower pH range 3.5 to 1.3, not only is ΔF^\ddagger for egg albumin about 700 cal. higher than for hemoglobin, but it is much more nearly constant, dropping by only 1000 cal. per pH unit. Apparently, if we were to speak, in terms of Steinhardt's interpretation, of a "real" ΔH^\ddagger for hemoglobin denaturation independent of pH, its value would not be greater than zero and might conceivably be negative.

As a résumé of the general denaturation behavior a few remarks may be made. The pH of maximum stability is on the same side of the isoelectric point as that for the maximum in the arithmetic sum of charges in the protein. In the case of the proteins this will be the pH of minimum entropy. Where the isoelectric point and the pH of maximum stability coincide, as in the case of hemoglobin, we expect that the basic and acidic bridge-forming groups are essentially equal in number. This need not necessarily mean exact equivalence in acid- and base-binding capacity, since at high pH the phenolic group of tyrosine, for example, may affect the titration curve but will not form a salt bridge.

When salt bridges are broken, we expect a certain binding or "freezing out" of polar solvent molecules on the freed charges (14c). Change of solvent from dilute aqueous solution to more concentrated solution will in every case affect this "freezing out" process. If the molecules frozen out are water, then addition of non-aqueous solvent molecules will lower the freezing point of the water, and less will be frozen out. We find that salt and alcohol both affect denaturation kinetics in the same way; both lower the freezing point of water and prevent its "freezing out." This effect is seen in the progressive increase of ΔS^\ddagger with increasing salt or alcohol concentration. The heat of freezing of this water is also lost, and ΔH^\ddagger is seen to go up with increasing salt and alcohol concentration.

Urea would, of course, have the same effect on the water molecules as would salt or alcohol, and yet its effect on the kinetics of denaturation is in the opposite direction. This probably means that the solvent frozen out is no longer water but urea, and that the bonding of urea to the activated complex is very strong. This is reflected in the large negative entropy of activation. In the absence of urea, at pH about 6, ΔS^\ddagger is about 300 units for the denaturation of egg albumin, whereas in concentrated urea solution at the same pH it is -90 units. There is thus a decrease of about 400 entropy units when water is replaced by urea as solvent. The net heat of binding the urea is also much greater than that for water. Thus ΔH^\ddagger for this same denaturation in water is about 135,000 cal., while in concentrated urea solution it is about -6000 cal., so that the net binding of urea is measured by a ΔH about 140,000 cal. more than is that of water.

This tendency of urea to form stronger bonds than water is to be ex-

pected when it is realized that the dipole moment of urea is about 8.6 Debye units (6) and that an 8 molal solution has a dielectric constant of about 108 at 0°C. (41).

A quantitative theory can be applied to take account of the change in activity coefficients with change in dielectric constant of the solvent. Solvent effects occur, however, which are independent of dielectric constant.

SOLVATION AND THE STRUCTURE OF PROTEINS

To make a hole in a liquid the size of the molecule requires the same amount of heat as to vaporize the molecule. To do both simultaneously requires twice the heat of vaporization. This fact is of interest in connection with theories of protein structure. Thus if one assumes space-enclosing structures such as those postulated by Wrinch (40), large quantities of solvent will be enclosed (1) in the interior of the molecule and (2) in the space between the protein molecules of a crystal. Spheres arranged in simple cubic structure have as much empty space between the spheres as the spheres themselves occupy. Actual structures which occur in crystals are never quite as loose as this. The unoccupied space for spherical molecules is ordinarily of the order of one-quarter of the occupied space. Although in ordinary crystals this space remains empty, protein molecules are so large that the solvent fills the lacuna. Steinhardt's (36) solubility results for pepsin give values which depend on the relative amounts of solid and solvent; this leads him to the conclusion that solid pepsin must be regarded as at least a two-component system. This would be expected if pepsin molecules were even approximately spherical.

Removal of the solvent from crystalline proteins by drying or by introducing it into a second solvent unable to replace the first should cause those molecules having large lacunas to collapse. The degree of instability is readily estimated from the above considerations if the size of the holes be known. Such a collapse of an enzyme would probably denature and inactivate it. Unfortunately the available experimental material does not justify a more detailed discussion at this time.

The authors wish to express their appreciation to Doctors M. L. Anson, M. Harris, and J. Steinhardt for helpful discussions.

REFERENCES

- (1) ANSON, M. L., AND MIRSKY, A. E.: *J. Gen. Physiol.* 17, 339 (1934).
- (2) ANSON, M. L., AND MIRSKY, A. E.: *J. Gen. Physiol.* 17, 393 (1934).
- (3) ANSON, M. L.: *Chemistry of the Amino Acids and Proteins*, edited by C. L. A. Schmidt, Chapter IX. Charles C. Thomas, Springfield, Illinois (1938).
- (4) ARRHENIUS, S.: *Immunochemistry*, p. 87. The Macmillan Company, New York (1907).

- (5) ASTBURY, W. T., AND WOODS, H. J.: *Phil. Trans. Roy. Soc.* **A232**, 333 (1933).
- (6) BERGMANN, E., AND WEIZMANN, A.: *Trans. Faraday Soc.* **34**, 783 (1938).
- (7) BOOTH, N.: *Biochem. J.* **24**, 1699 (1930).
- (8) BURK, N. F.: *J. Biol. Chem.* **120**, 63 (1937).
- (9) CHICK, H., AND MARTIN, C. J.: *J. Physiol.* **43**, 1 (1911).
- (10) COHN, E. J., McMEKEN, F. L., EDSELL, J. T., AND BLANCHARD, M. H.: *J. Biol. Chem.* **100**, No. 3, xxviii (May, 1933).
- (11) CUBIN, H. K.: *Biochem. J.* **23**, 25 (1929).
- (12) ESCOLME, A. I., AND McC. LEWIS, W. C.: *Trans. Faraday Soc.* **23**, 651 (1927).
- (13) v. EULER, H., AND LAUREN, I.: *Z. physiol. Chem.* **108**, 64 (1919).
- (14) (a) EYRING, H.: *J. Chem. Phys.* **3**, 107 (1935).
 (b) WYNNE-JONES, W. F. K., AND EYRING, H.: *J. Chem. Phys.* **3**, 492 (1935).
 (c) STEARN, A. E., AND EYRING, H.: *J. Chem. Phys.* **5**, 113 (1937).
 (d) STEARN, A. E.: *Ergeb. Enzymforsch.* **7**, 1 (1938).
- (15) FAMULENER, L. W., AND MADSEN, T.: *Biochem. Z.* **11**, 186 (1908).
- (16) GERLOUGH, T. D., AND BATES, R. W.: *J. Pharmacol.* **45**, 19 (1932).
- (17) HOPKINS, G. F.: *Nature* **126**, 328, 383 (1930).
- (18) LA MER, V. K.: *Science* **86**, 614 (1937).
- (19) LAVINE, T. F.: *J. Biol. Chem.* **117**, 309 (1937).
- (20) See LEWIS AND RANDALL: *Thermodynamics*, p. 337. McGraw-Hill Book Company, Inc., New York (1923).
- (21) LEWIS, P. S.: *Biochem. J.* **20**, 965, 978 (1926).
- (22) LEWIS, P. S.: *Biochem. J.* **20**, 984 (1926).
- (23) LUERS, H., AND LANDAUER, M.: *Z. angew. Chem.* **35**, 469 (1922).
- (24) LUERS, H., AND WASMUND, W.: *Fermentforschung* **5**, 169 (1922).
- (25) MCGILLIVRAY, I. H.: *Biochem. J.* **24**, 891 (1930).
- (26) MIRSKY, A. E., AND PAULING, L.: *Proc. Natl. Acad. Sci. U. S.* **22**, 439 (1936).
- (27) NICLOUX, M.: *Compt. rend. soc. biol.* **56**, 839 (1904).
- (28) NORTHRUP, J. H.: *Ergeb. Enzymforsch.* **1**, 302 (1932).
- (29) PACE, J.: *Biochem. J.* **24**, 606 (1930).
- (30) PACE, J.: *Biochem. J.* **25**, 1 (1931).
- (31) PACE, J.: *Biochem. J.* **25**, 442 (1931).
- (32) PACE, J.: *Biochem. J.* **25**, 1485 (1931).
- (33) SHINOHARA, K., AND KILPATRICK, M.: *J. Biol. Chem.* **105**, 241 (1934).
- (34) SOOKNE, A. M., AND HARRIS, M.: *J. Research Natl. Bur. Standards* **19**, R. P. 1043 (1937).
- (35) STEINHARDT, J.: *Kgl. Danske Videnskab. Selskab, Math.-fys. Medd.* **14**, No. 11 (1937).
- (36) STEINHARDT, J.: *Proceedings of the Thirty-second Annual Meeting of the American Biological Chemists*, March 30, 1938.
- (37) (a) SPEAKMAN, J. B. AND STOTT, E.: *Trans. Faraday Soc.* **30**, 539 (1934).
 (b) SPEAKMAN, J. B.: *Textile Mfr.* **62**, 236 (1936).
 (c) SPEAKMAN, J. B.: *Textile Recorder* **54**, 36 (1936).
 (d) SPEAKMAN, J. B., AND HIRST, M. C.: *Nature* **128**, 1073 (1931).
 (e) SPEAKMAN, J. B., AND HIRST, M. C.: *Trans. Faraday Soc.* **29**, 148 (1933).
 (f) SPEAKMAN, J. B., AND TOWNSEND, F.: *Trans. Faraday Soc.* **32**, 897 (1936).
- (38) SVEDBERG, T.: *Trans. Faraday Soc.* **26**, 740 (1930).
- (39) TAMMAN, G.: *Z. physik. Chem.* **18**, 426 (1895).
- (40) WRINCH, D. M.: *Phil. Mag.* [7] **24** (Supplement), 940 (1937).
- (41) WYMAN, J., JR.: *J. Am. Chem. Soc.* **55**, 4116 (1933).
- (42) ZILVA, S. S.: *Biochem. J.* **8**, 656 (1914).

ELECTROPHORESIS OF PROTEINS BY THE TISELIUS METHOD¹

L. G. LONGSWORTH AND D. A. MACINNES

Laboratories of The Rockefeller Institute for Medical Research, New York, New York

Received January 30, 1939

INTRODUCTION

The recent improvements introduced by Tiselius (15) in the electrophoretic method have greatly increased its importance in the study and preparation of proteins. Our contribution to this Symposium will consist of a discussion of the principles underlying the Tiselius method, together with a description of the experimental procedure involved in the use of the method, including such modifications as we have introduced. The topics to be considered are (a) the apparatus, (b) the optical principles involved in the observation of electrophoretic boundaries, (c) the effects of thermal convection, (d) some typical applications of the method, and (e) "boundary anomalies". These will be treated in the order given.

APPARATUS

To study electrophoresis of proteins and related materials by the moving boundary method it is necessary, in the first place, to form a boundary between a solution of the material in a suitable buffer and the buffer itself. Passage of current then causes the boundary to move. For this movement to be a measure of the mobility of the protein it is desirable, first, that the electric field and pH in the neighborhood of the boundary be substantially constant, second, that the electrode processes do not involve the evolution of gas or other uncertain volume changes, and third, that the electrode products do not reach the regions in which the boundaries are moving. These conditions have been admirably met in the electrophoretic apparatus developed by Tiselius, which, with the modifications that we have made, will be described below.

The cell in which the boundaries are formed and observed is shown in cross section in figure 1A and consists of the sections I, II, III, and IV.

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

These may be slid over one another along the planes a-a', b-b', and c-c'. Through the cell runs a U-shaped channel d-d' of rectangular cross section. Figure 1B is a top view of one of the center sections. To form a boundary, the channel is filled with the buffer solution of protein to a level slightly above the plane b-b', about 10 ml. of solution being required. Section III is then pushed to one side, the excess solution in section II is removed, and this section is rinsed with buffer. The remainder of the cell and the attached electrode vessels to be described are then filled with the buffer.

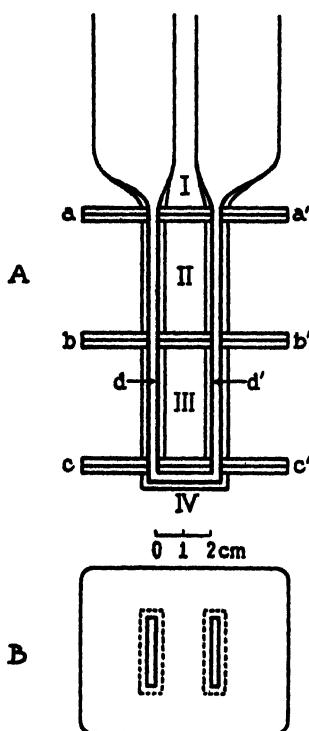


FIG. 1. A, electrophoresis cells in cross section; B, top view of one of the center sections.

The support for the cell and the electrode vessels, together with the mechanism for moving the sections of the cell in relation to each other, are shown in figure 2. In our apparatus a rack-and-pinion system replaces the pumps utilized by Tiselius for this purpose. Turning the knurled knobs k-k' attached to the concentric shafts (s) operates the bevel gears (g), which in turn cause a horizontal motion of the racks r. Each rack presses against a metal insert e, which communicates the pressure

to the edges of the horizontal glass plates. The sections II and III may therefore be shifted in either direction by manipulation of the appropriate knurled knob.

After filling the apparatus as described above, the silver-silver chloride electrodes (E-E' of figure 2) are inserted. For effective operation of these electrodes, described more fully later, they must be immersed in a strong chloride solution. This is accomplished by carefully introducing the solu-

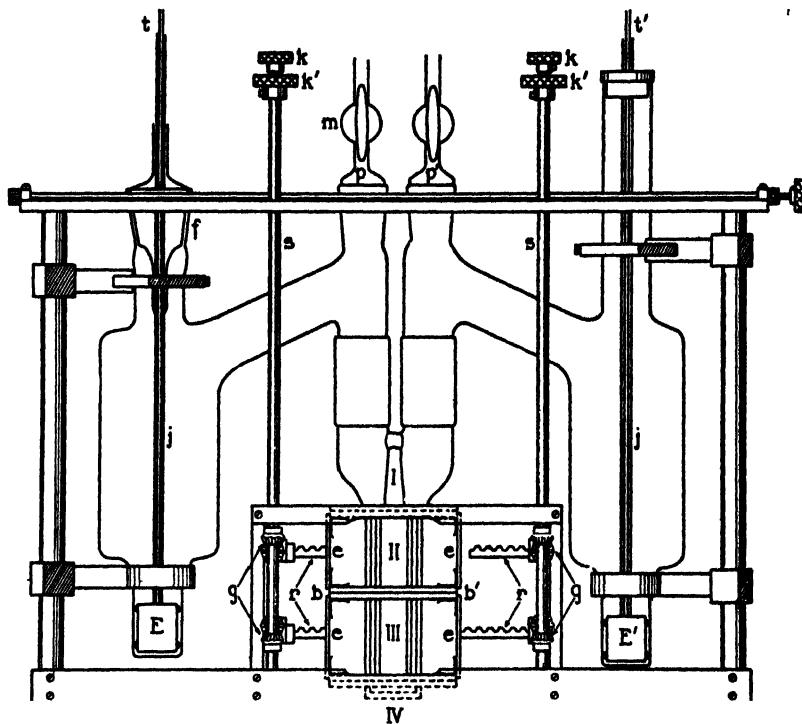


FIG. 2. Electrophoresis cell, electrode vessels, and support

tion through the silver tubes t-t'. These silver tubes, which are insulated by the glass tubes (j), also serve as current leads to the electrodes E-E'. We have modified the procedure of Tiselius and our earlier procedure, in that one side of the apparatus is closed, care being taken to exclude air bubbles. This is accomplished with the ground-glass stopper f and the stopcock m. Closing one side is more convenient than having both sides open, as the latter procedure involves equalizing the liquid levels on the

two sides of the system before forming the boundaries. It also permits the use of an improved type of compensation, to be described below.

With the cell and electrode vessels filled as described, the two boundaries are formed in the plane b-b' by returning section III to the position shown. A potential from a battery or "power pack" applied to the terminals t-t' will then, in general, cause the boundary in one side of the cell to rise and that in the other to fall.

With the design of the apparatus shown in figure 2, filling of the cell and electrode vessels, as described, may be carried out in the low-tem-

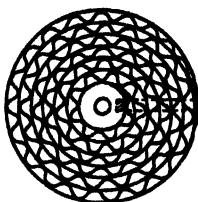
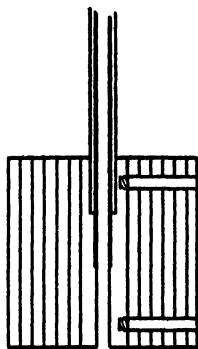


FIG. 3. Diagram illustrating the electrode construction

perature thermostat. Since the stoppers p and p' are directly above the cell, the protein solution and buffer may be introduced with pipets. For this purpose we have found a syringe, provided with a long stainless-steel needle, serviceable.

Further details of the electrodes E-E' are as follows. They must be capable of carrying currents of at least 30 milliamperes for long periods of time without evolution of gas. Sufficient capacity has been obtained by winding a flat and a corrugated strip of sheet silver together into a tight spiral, as shown in figure 3. The ends of the spirals are anchored to a hollow silver core with silver screws. The silver tube is also threaded

into this core. This type of construction, suggested by Professor H. P. Cady, exposes a large electrode surface to the electrolyte. Air bubbles trapped by the electrodes may be easily dislodged.

OBSERVATION OF THE BOUNDARIES

Another important contribution to the electrophoretic method made by Tiselius was the adaptation of Toepler's "schlieren" (shadow) method (9) for the observation of the boundaries. A diagram of the optical system is shown in figure 4. The image of the slit S, illuminated by the lamp L and condenser C, is brought to focus in the plane P by the lens D. The schlieren diaphragm, a screen with a sharp, horizontal upper edge, is placed in the plane P and may be moved vertically, a micrometer adjustment being used. The cell E, in which the electrophoresis is carried out, is placed as near the lens D as the thermostat construction permits. The

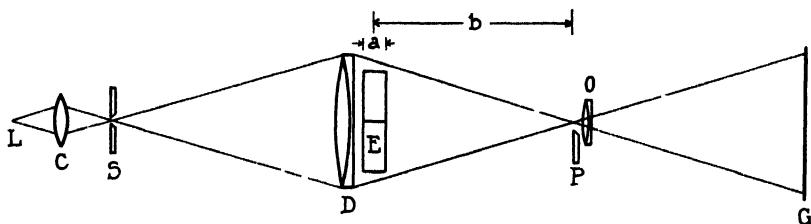


FIG. 4. Arrangement for observation of the boundaries with the schlieren method L, 5-volt, 18-ampere ribbon filament lamp; C, projection condenser; S, 0.3 × 15 mm horizontal slit, D, Dallmeyer portrait lens 6 D, 4 in. diameter, 24 in. focal length, E, electrophoresis cell; P, schlieren diaphragm, G, ground-glass screen or photographic plate; O, Dallmeyer R R lens, 3 in. diameter, 24 in. focal length

camera objective O, placed immediately behind the schlieren diaphragm, is focussed on the cell and forms a full-size image on a ground-glass or photographic plate at G. Further details of the system as used in obtaining the results described in this paper are indicated in the legend of the figure.

In the absence of refraction gradients in the electrophoresis cell all of the light is brought to focus in the image of the illuminated slit at P and enters the camera objective. If, however, a boundary is present in the cell, the refraction decreases with increasing height through the boundary, and the pencils of light through this region are deflected downward. These deflected pencils are intercepted by the schlieren diaphragm and fail to reach the screen. Thus the region at G conjugate to the boundary appears as a dark band on a light background. This is shown in figure 5. The horizontal dark lines are the schlieren bands of the boundaries between a

0.5 per cent egg albumin solution in a 0.02 normal sodium acetate buffer at pH 5.2 and the pure buffer. The upper and lower photographs are of the boundaries migrating in the anode and cathode sides of the channel, respectively. Exposures were made at 30-min. intervals, and after four exposures the current was reversed. It is of great interest that the reversal brought the boundaries accurately back to their original positions.

Some of the characteristic features of the schlieren method may be considered with the help of figure 4. The angular deviation of a pencil of light in the boundary is proportional, under appropriate conditions, to

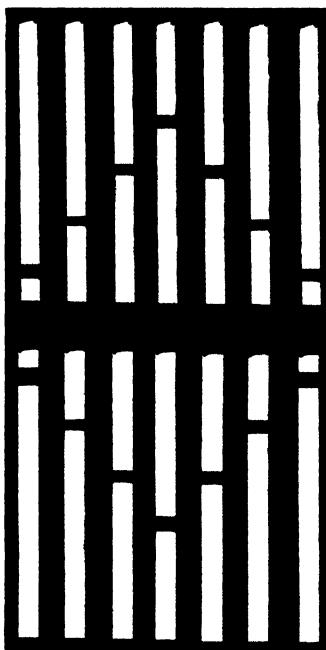


FIG. 5. Schlieren bands of a single protein. Exposures were made at 30-min. intervals and the current was reversed after the fourth exposure.

the gradient, dn/dx , of the refractive index, n , and the horizontal breadth, a , of the boundary. The displacement, Δ , of the schlieren diaphragm necessary to intercept the deflected pencil is also proportional to the "optical lever arm," b . We therefore have the expression

$$\Delta = ab \frac{dn}{dx} \quad (1)$$

in which a and b are constants of the apparatus, and dn/dx varies vertically through the boundary. As the schlieren diaphragm is raised, the first pencils of light to be intercepted are those which have passed through

the steepest gradients of refractive index, i.e., the center of the boundary if the diffusion and spreading of the latter have been normal. Thus a series of photographs of the boundary with decreasing displacement of the diaphragm give an indication of the variation of the refractive index through the boundary. Such a series of photographs of the boundaries formed in a mixture of rabbit and guinea pig hemoglobins is shown in figure 6, the upper half of the figure referring to rising boundaries and the lower to descending boundaries.

Since the refraction of the solution is substantially proportional to the concentration of protein, the schlieren method may be used to obtain

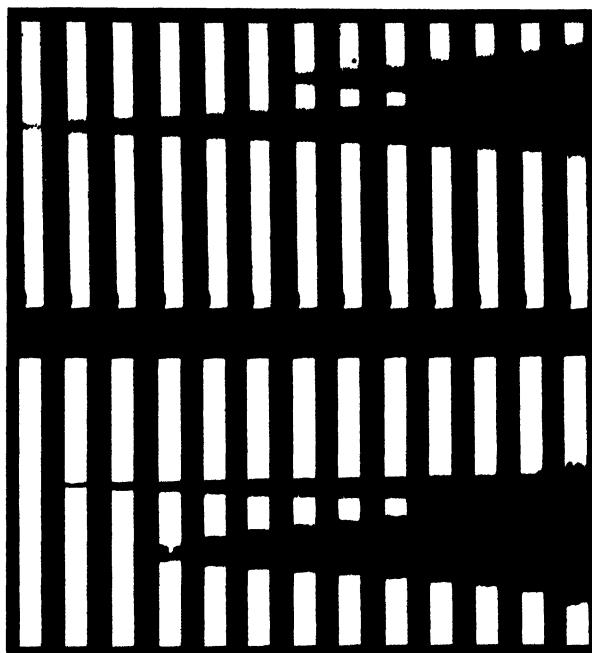


FIG. 6. Schlieren bands, for different schlieren diaphragm settings, of a mixture of rabbit and guinea pig hemoglobins.

quantitative estimates of the concentrations of the various constituents. Thus, if the cell image at the photographic plate is masked by a narrow vertical slit, a slow horizontal movement of the plate simultaneously with the vertical displacement of the schlieren diaphragm produces automatically on the plate a transparent area proportional to $\int \Delta dx$. Reference to equation 1 shows that

$$\Delta dx = \int ab \frac{dn}{dx} dx = ab\Delta n$$

in which Δn is the refractive index increment due to the protein constituent causing the boundary in question. This modification of the schlieren method for quantitative analysis has been described in two papers from this laboratory (6, 7).

REDUCTION OF THERMAL CONVECTION

The chief problem in electrophoretic methods is the elimination of mixing due to thermal convection. If a current of electricity is passing through a conducting solution in a tube, heat is generated in each volume element of the solution but flows to the thermostat only through the wall of the tube. The solution along the axis of the tube is thus hotter than at the wall. Normally, therefore, the solution at the wall will be heavier, and in falling will give rise to convection currents. We shall consider, as a typical example, a current of 0.006 ampere passing through a solution the specific conductance of which is 0.0038 mhos (0.1 N sodium acetate at 0°C.) in a cylindrical tube of 5 mm. internal and 10 mm. external diameter. The formula² which describes the temperature of the solution, t_s , in the steady state as a function of the distance, r , from the axis is

$$t_{r=0} - t_s = \frac{I^2 \epsilon}{4\kappa K_s} r^2 \quad (2)$$

The corresponding formula for the temperature of the glass, t_g , is

$$t_g - t_0 = \frac{a^2 I^2 \epsilon}{2\kappa K_g} \ln \frac{b}{r} \quad (3)$$

In these equations a and b are the inside and outside radii, respectively, of the tube, K , and K_g are the thermal conductivities of the solution and glass, I is the current density, ϵ is the electrical equivalent of heat, and t_0 is the thermostat temperature. Using equation 2, the solution along the axis of the tube is 0.65°C. hotter than at the wall, and from equation 3 the drop in the wall is 0.67°C. The computed temperature distribution is given in figure 7a, in which the temperature increase, Δt , over that of the thermostat is plotted against the distance from the axis of the tube. If the thermostat is regulating at 25°C., this temperature gradient in the buffer solution is accompanied by the density variations shown in figure 7b and, as has been stated, it is these differences which cause mixing, by convection currents, of the solution in the tube. If, on the other hand,

² These equations are from a private communication from Dr. Melvin Mooney. The authors wish to acknowledge gratefully this aid, which has been of service to them both in this work and in their earlier work on the determination of transference numbers by the moving boundary method.

the thermostat is regulating at 0°C., the density differences in the solution are much less and the variation is in the opposite direction, as shown in figure 7c. The contrast between the curves of 7b and 7c arises from the fact that this buffer solution has a maximum density at 2.85°C. (2). If, in this particular example, the thermostat temperature were regulated at 1.85°C., the average temperature in the tube would be 2.85°C., and the density gradient would be a minimum, as indicated by the horizontal

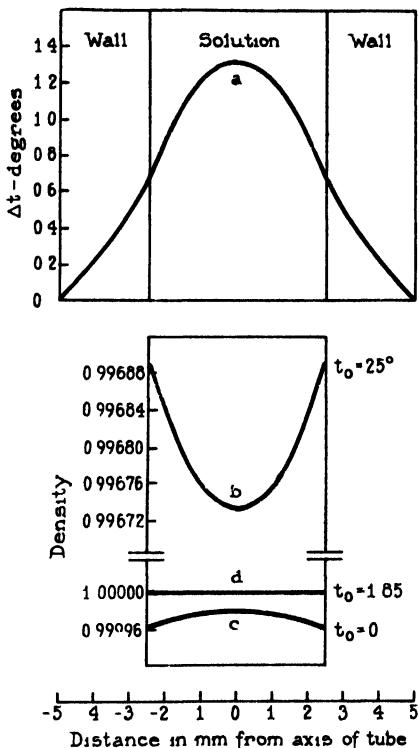


FIG 7 Distribution of temperature and density in a salt solution, in a cylindrical tube, during passage of electric current

line in figure 7d. Some preliminary measurements of the temperature variations in a rectangular channel indicate that they are of the same order of magnitude as those indicated in the example just given for a cylindrical tube.

In spite of the fact that the density differences produced by temperature gradients are, as indicated in figure 7, very small, the recent great advances in the electrophoretic method are due to the important observation of

Tiselius that convection can be largely eliminated by working at temperatures near that of maximum density. The computations suggest that further progress may be made by additional knowledge of the temperature of maximum density of the buffer solutions used and of the actual temperature distribution in the electrophoresis cell. Investigations along both these lines are in progress by Dr. T. Shedlovsky of this laboratory.

SOME TYPICAL APPLICATIONS OF THE ELECTROPHORETIC METHOD

One of the most interesting uses of this electrophoretic method is the determination of the mobilities of proteins as a function of the pH and the ionic strength of the buffer solution in which they are dissolved. The following data for a typical experiment were obtained in collaboration with Drs. Karl Landsteiner and van der Scheer (4). The protein was crystalline albumin from guinea hen eggs. A 0.5 per cent solution was prepared in a 0.02 normal acetate buffer at pH 5.19. The specific conductances, at 0°C., of the buffer and protein solutions, as measured on the bridge described by Shedlovsky (10), were essentially the same and equal to 0.000832 mho. The current, measured potentiometrically, was practically constant throughout the experiment and equal to 0.00679 ampere. From a calibration with mercury the cross-sectional area of the channel in the cell was known to be uniform and equal to 0.815 sq.cm. The current density was, therefore, 0.00833 ampere and the electric field 10.01 volts per centimeter. The schlieren bands were photographed, using unit magnification, at 30-min. intervals and are shown in figure 5. As mentioned earlier in connection with this figure, the current was reversed after the fourth exposure and the boundaries returned accurately to their original position. The displacement of the boundary moving upward into buffer during the 90-min. interval before reversal of the current was 2.05_b cm., giving $3.80_b \times 10^{-4}$ cm. per second as the velocity and -3.80×10^{-5} cm. per second per volt per centimeter as the mobility of the protein, the migration being anodic. The corresponding displacement of the boundary moving into protein was 1.99_b cm., giving a mobility of -3.70×10^{-5} . The average from the two boundaries is thus -3.75×10^{-5} . The greater displacement of the boundary moving into buffer has been quite generally observed and will be discussed later in this paper.

In addition to the measurement of mobilities the method furnishes information as to the purity and homogeneity of the protein solution and also may be adapted to separation of the components of a mixture. This latter application, called "electrophoretic analysis" by Tiselius, is illustrated in figure 8. Suppose a mixture of proteins A, B, and C, for which the (positive) mobilities are $u_A > u_B > u_C$, is placed in the cell as shown in figure 8a. Movement of the bottom center section of the cell to the

right will bring the protein and buffer solutions into contact in the plane α . On passage of a current three boundaries will appear in the upper cathode and in the lower anode sections, as is shown in figure 8b. It is evident from the figure that, in the ideal case assumed, there is a separation of pure component A in the region between the two leading boundaries on the cathode side and of pure component C between the two slowest boundaries on the anode side. However, if the electrolysis is continued as indicated, before any large proportions of A and C have been separated, the boundaries will have migrated out of the cell in one case and into the bottom section in the other. However, if the second boundary could be given an apparent velocity of zero, as indicated in figure 8c, while the leading boundary moves through the length of one section, isolation of this section

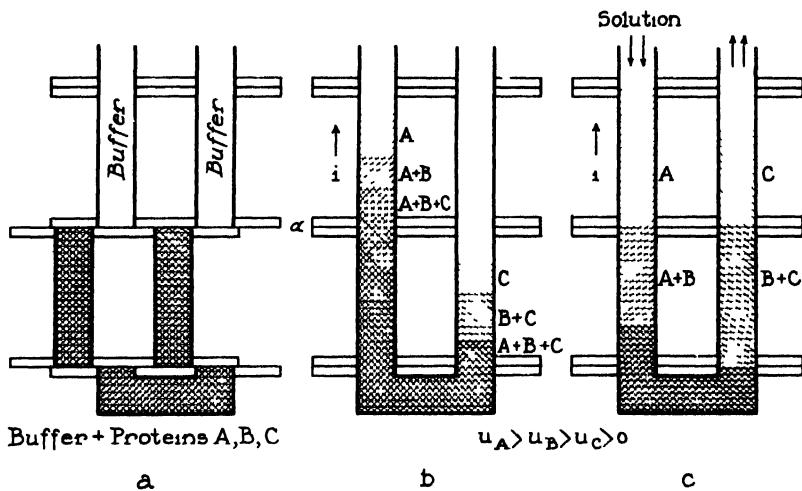


FIG. 8. Ideal electrophoresis of a protein mixture

from the others would make possible the recovery of a solution of pure A from the upper cathode portion of the cell. Simultaneously, as indicated, the slowest component would have an apparent negative velocity, and pure C could be recovered from the upper anode section. As a matter of fact, a boundary can be given any apparent velocity desired by a displacement of the entire solution of the cell. Such displacements have been accomplished in a variety of ways. In Smith's (11) moving boundary apparatus this was done by the withdrawal of mercury, while Collie and Hartley (1) used a piston driven by clockwork. In Tiselius' apparatus he displaces the boundaries by withdrawing, with clockwork, a loosely fitting plunger in one of the electrode vessels. We have used this method also, but in the more recent apparatus described in this paper we have

kept one electrode vessel closed and have displaced the solution in the cell by forcing buffer into this side from a syringe, the piston of which is displaced, at the desired rate, by a threaded rod operated by a synchronous clock motor.

Figure 9 is the electrophoretic pattern, obtained in collaboration with Dr. Karl Landsteiner, of a water-insoluble fraction of a naturally occurring mixture of plant proteins. The three bands indicate the presence of three well-defined constituents. Having demonstrated the complexity of the protein material, it was of importance to locate a certain biological activity which one of the components possessed. To this end the leading components were separated, using the compensation method outlined



FIG. 9

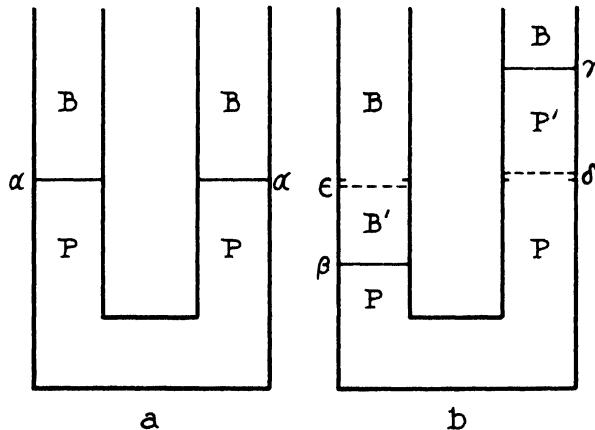


FIG. 10

FIG. 9 Electrophoretic pattern of water-insoluble fraction of a mixture of plant proteins

FIG. 10 Diagrammatic representation of boundary formation and migration during electrophoresis

above. They exhibited no activity, whereas the slow component did possess the activity. An electrophoretic study of the water-soluble fraction, which was also active, indicated the presence of a single protein with a mobility near that of the active material in the water-insoluble fraction. The possibility remained that the two active proteins were identical, the ammonium sulfate separation having possibly been incomplete. This question was answered unambiguously by analyzing electrophoretically a mixture of the two active proteins. Two bands were obtained, indicating the separate identity of the two proteins.

In electrophoretic analysis, using compensation, it is frequently necessary to pass relatively heavy currents through the cell for long periods of

time. Under these conditions either very large electrode vessels must be used, or a layer of concentrated buffer solution must be introduced (17) between the concentrated chloride solution around the electrode and the dilute buffer solution above. Tiselius (15) has given an excellent discussion of the size of electrode vessel required when no concentrated buffer solution is used.

BOUNDARY ANOMALIES

An electrophoresis experiment with a single protein may be represented diagrammatically by figure 10. Initially boundaries are formed at the level α in the two sides of the cell between the buffer solution at a concentration B and the buffer solution of protein at a concentration P (in which the protein is assumed to be negatively charged). Passage of a current causes one of these boundaries to descend through a volume V_P to a new position β (figure 10b), and the other to rise through a volume V_B to the position γ . Under ideal conditions, realized closely with a dilute solution of a homogeneous protein in a buffer of sufficient strength to control the pH and conductance of the system, the volumes V_P and V_B will be equal. With more concentrated protein solutions, however, so-called "boundary anomalies" are observed. These are illustrated in figures 11 to 13. For instance, the schlieren bands shown in figure 11a for different settings of the schlieren diaphragm indicate a boundary that has risen into buffer and 11b the boundary that has descended into the protein solution in the other side of the cell. It is quite evident that the boundary moving into buffer is the sharper of the two. This is quite generally observed. A more disturbing anomaly is shown in figure 12, in which the relative displacement of the schlieren bands indicates that the rising boundary is moving more rapidly than the descending one; this is also, apparently, a general rule. Another anomaly is illustrated in figure 13. In this experiment with a 2.5 per cent solution of a single protein the boundaries were formed in the usual manner. They were then shifted in the cell with the compensation device to the positions shown in column 1, so that a very slowly moving boundary, which might normally be obscured by the horizontal glass plates, could become visible. On passage of a current the protein boundaries migrated as usual to the new positions indicated in column 2 of the figure. However, on raising the schlieren diaphragm a second, and relatively faint, boundary, δ (column 3), appeared in the protein solution but slightly removed from the original position of the boundary. In the discussion to follow it will become evident that this is not due to a second protein, but arises from a gradient of concentration left behind by the advancing protein boundary. It is similar to the " δ globulin" boundary of Tiselius (16, 13). On raising the diaphragm still further, at a

somewhat later time, a very faint boundary appeared in the buffer solution at ϵ (column 4) and was similar in nature to the δ -boundary.

In practice the protein solution is prepared by dialysis against the buffer solution. It is important to recall that a difference of buffer salt concentration, between the two solutions, exists when the dialysis is complete,



FIG. 11

a
b

FIG. 12

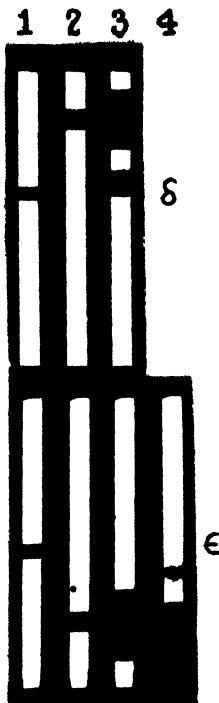


FIG. 13

 ϵ

FIG. 11 Schlieren bands, for different diaphragm settings, of a single protein. The boundary (a) moving into buffer is sharper than that (b) moving into protein.

FIG. 12. Schlieren bands of a single protein. The velocity of the boundary moving into buffer is greater than that of the boundary moving into protein.

FIG. 13. Schlieren bands of a single protein. The δ and ϵ bands are due to different concentrations of the same substance.

owing to the Donnan equilibrium. When passage of a current causes a boundary between two such solutions to move (from α to β , figure 10) there is formed in the intervening volume, V_P , a buffer solution of composition B' . This composition has been "adjusted," in general, to a value different from B in such a way that its "regulating function" has the same value as that of the protein solution it has replaced. To quote an earlier

paper by the authors (8), "This (regulating) function defines a property of the solution which, at any given point, retains a constant value independent of changes of concentration caused by electrolytic migration. If, as a result of such migration, species of ions different from those initially present appear at a point, their concentrations will be adjusted to values compatible with the constant determined by the *initial* composition of the solution." It may be noted here that the concentration changes which occur behind the boundary moving into the protein solution are

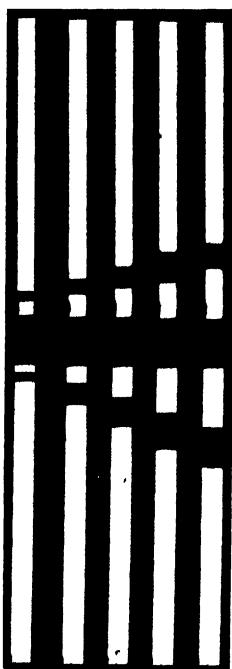


FIG. 14. Schlieren bands of the concentration boundaries between 0.01 and 0.1 normal solutions of lithium chloride.

similar to those encountered by one of the authors (5) in the electrolysis of mixtures of hydrogen and potassium chlorides. The boundary is thus formed between two solutions of the same salt, but at the different concentrations B and B' . It is stable if the lower solution is the more concentrated. Boundaries of this type have been investigated theoretically by Kohlrausch (3) and Weber (18) and experimentally by Smith (12) and ourselves. That such boundaries persist and move on the passage of a current is illustrated in figure 14. Boundaries were formed between 0.01 and 0.1 normal solutions of lithium chloride; these were shifted from behind

the horizontal plates and current was passed. The motion of the boundaries, actually very slow compared with others dealt with in this paper, is indicated by the displacement of the schlieren bands. The motion is due to a change of the transference numbers with concentration.

With the boundary moving into buffer there is a similar but more complicated adjustment of the composition of the protein solution which replaces the buffer as the boundary rises. The resulting concentration boundary, δ (figure 10), between the solutions P and P' moves slowly under the influence of the current and, as mentioned above, is similar to the " δ globulin" boundary of Tiselius. The greater visibility of this boundary, compared with the ϵ -boundary, may be due to the fact that the first involves a gradient of protein concentration whereas the second does not.

In order to account for the boundary anomaly illustrated in figure 11, i.e., the greater diffuseness of the boundary moving into the protein solution, we must consider the effect of the concentration changes, discussed in the preceding paragraphs, upon the specific conductances of the solutions. Owing mainly to its relatively high viscosity, the specific conductance, κ_P , of the protein solution is generally lower than that, κ_B , of the buffer. The conductance differences at the actual boundaries, β and γ of figure 10b, are further increased by the fact that $\kappa_{B'} > \kappa_B$ and $\kappa_{P'} < \kappa_P$. Consequently for a given current the electric field is greater in the protein solution than in the buffer, and variations of the field exist at the boundaries. The dilute uppermost layers of the boundary β moving into the protein solution thus find themselves in weaker fields than do the more concentrated layers and thus tend to lag behind, causing the boundary to become diffuse. In the case of the boundary γ moving into buffer, however, the dilute, slowly moving layers tend to be overtaken by the faster, concentrated ones, with the result that the boundary tends to become sharper. (See also reference 14, page 28.)

If there is a stable δ -boundary, as shown in figure 13, the total concentration below the boundary must be greater than above it. Moreover, it was found in an actual case that the solution P' had a lower conductance than the solution P . The potential gradient above the δ -boundary will therefore, in general, be greater than below, with the result that the protein must move more rapidly in that region than in the main body of the solution. If the displacement of the boundary γ can be assumed to be a measure of the velocity of the protein particles in the solution P' , this boundary would be expected to move more rapidly than the β -boundary, as has been true in every case we have observed.

In the preceding discussion of boundary anomalies we have ignored the small differences of pH, required by the Donnan equilibrium, which exist between the buffer and protein solutions. Owing to the change of protein

mobility with pH, these differences at the boundaries can cause anomalies similar to those mentioned. Both theory and experiment indicate, however, that by far the major portion of these anomalies arises from the differences in specific conductance.

Although boundary anomalies are always present to some extent they may be reduced by proper selection of the buffer solution. The buffer used as solvent for the protein should have a high capacity in order to reduce, relatively, the buffer action of the protein itself. It should also have a low specific conductance in order to decrease disturbances due to the heating effect of the current. Since both buffer capacity and conductance increase with the concentration of buffer salts, it is evident that these two conditions are mutually incompatible, and a compromise must be made. As buffer capacity does not depend upon ionic mobilities, buffer salts the ions of which have low mobilities should be selected if possible. We have used sodium salts in preference to potassium salts for this reason. Lithium salts would be still better, if circumstances justified their preparation. High concentrations of weak electrolytes should be avoided in the preparation of buffers, unless allowance is made in the thermostat temperature for the large effect of these components on the temperature of maximum density of the solution.

REFERENCES

- (1) COLLIE AND HARTLEY: Trans. Faraday Soc. **30**, 657 (1934).
- (2) International Critical Tables, Vol. III, p. 108 McGraw-Hill Book Company, Inc., New York (1929).
- (3) KOHLRAUSCH: Ann. Physik **62**, 209 (1897).
- (4) LANDSTEINER, LONGSWORTH, AND VAN DER SCHEER: Science **88**, 83 (1938).
- (5) LONGSWORTH: J. Am. Chem. Soc. **52**, 1897 (1930).
- (6) LONGSWORTH: J. Am. Chem. Soc. **61**, 529 (1939).
- (7) LONGSWORTH: Annals N. Y. Acad. Sci., in press (1939).
- (8) MACINNES AND LONGSWORTH: Chem. Rev. **11**, 171 (1932).
- (9) SCHARDIN: Das Toeplersche Schlierenverfahren Grundlagen für seine Anwendung und quantitative Anwendung VDI-Verlag G.m.b.H., Berlin (1934).
- (10) SHEDLOVSKY: J. Am. Chem. Soc. **52**, 1793 (1930).
- (11) SMITH: J. Am. Chem. Soc. **50**, 1904 (1928).
- (12) SMITH: J. Research Natl. Bur. Standards **6**, 917 (1931).
- (13) STENHAGEN: Biochem. J. **32**, 714 (1938).
- (14) TISELIUS: Thesis, Upsala, 1930. Nova Acta Regiae Soc. Sci. Upsaliensis **4**, 7.
- (15) TISELIUS: Trans. Faraday Soc. **33**, 524 (1937).
- (16) TISELIUS: Biochem. J. **31**, 1464 (1937).
- (17) TISELIUS: Svensk Kem. Tid. **50**, 58 (1938).
- (18) WEBER: Die partiellen Differential-Gleichungen der mathematischen Physik, 5th Auflage 1, 195 203, 503-27. Braunschweig (1910).

OPAQUE OR ANALYTICAL ULTRACENTRIFUGES¹

JAMES W. McBAIN

Department of Chemistry, Stanford University, California

Received January 30, 1939

When sand or gravel is dropped into quiet water, it falls at a definite rate which reflects its size or weight of particle. If there are several sizes, each has its corresponding rate. The same is true for colloidal particles and ordinary molecules, if the water is placed in a sufficiently intense field of force and yet held quietly free from vibration and stirring. This is the principle of the ultracentrifuge.

DEFINITIONS AND PURPOSES

Ultracentrifuges are centrifuges of low or high power containing a fluid in which convection is avoided, so that molecules or particles may sediment and diffuse undisturbed. In transparent ultracentrifuges these processes may be followed by measuring the absorption of light or the index of refraction.

In analytical ultracentrifuges all the expensive optical accessories required for all forms of transparent ultracentrifuges are dispensed with. Instead, at any given time, a whole or a portion of the contents of the ultracentrifuge is removed for analysis by any suitable chemical, physical, or biological method. This analysis determines what movement of each component of the fluid has occurred through a definite cross section of definite radius in a definite field of force.

Velocity of sedimentation, s , is customarily expressed as the linear radial rate of movement in centimeters per second per dyne of centrifugal force. It is an arbitrary convention to record as s_{20} the observed value of s multiplied by the ratio between the measured viscosity of the actual system and that of water at 20°C., correcting the density of the solution to the value for pure water at 20°C.

The sedimentation velocity of ordinary molecules requires such high centrifugal fields that it is at present beyond the range of any existing

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

transparent ultracentrifuge. As will be seen, however, it should lie within the scope of the high-power analytical ultracentrifuge.² Sedimentation velocity is measured in the highest practical field with the dual purpose of rendering it so rapid as to be undistorted by diffusion and of discriminating more sharply between different particles present. Direct analytical measurement of each moving component may serve to determine the composition of a particle and also to show which components move or vary independently. The dependence of the rate upon various factors yields such information as that of association or of the splitting or dissociation of proteins in solvents of exceptionally high dielectric constant (3, 4, 32, 34, 39, 40) or with certain added agents. With certain restrictions it may measure the density and hydration of the particles or molecules (18, 15, 7).

Sedimentation equilibrium is measured in comparatively low fields, except for the smallest molecules. It is the final stationary state where, at any point, velocity of sedimentation is exactly balanced by diffusion. Its advantage is its thermodynamic relation to molecular or particle weight independent of shape although subject to the usual restrictions as to activity coefficient. In practice this has been set equal to unity. The effect of extreme length and flexibility of a molecule, which distorts ordinary molecular weight determinations (26), has not yet been sufficiently examined in the ultracentrifuge. In mixtures and polydisperse systems each component attains its own dynamic equilibrium. In the opaque ultracentrifuge one analysis at one speed is enough to measure the equilibrium, provided the analysis singles out a monodisperse constituent. Otherwise, runs at different speeds and preferably with analysis of more than one portion of the sedimented column are required.

THEORY

In the analytical ultracentrifuge the one factor, sedimentation velocity, may be isolated for exact measurement, uninfluenced by the other factor, diffusion. In contrast to the transparent ultracentrifuge it is a matter of complete indifference as to whether or not there is a moving boundary somewhere in the immobilized portion. The precision of the result depends upon the quantitative analysis, the measurement of the radius, temperature, and speed, together with the precautions taken in design

² *Footnote added in proof:* Mr. F. A. Leyda in the author's laboratory has measured the sedimentation velocity of sucrose, using Pliofilm 0.0015 in. thick for spacing discs in the insert and using only 2110 r.r.s. with the 37-mm. rotor. He obtained $s_{20} \times 10^{12} = 0.170$ and 0.174. When these observations were corrected for diffusion, they became 0.21, whereas the calculated theoretical result for anhydrous sucrose is 0.23. The lower value is reasonably ascribable to hydration.

and manipulation to eliminate the disturbing effects of convection and diffusion. The analytical method is therefore potentially the most accurate and is still applicable where, owing to the absence of a boundary, the transparent ultracentrifuge fails.

Mason and Weaver (17, 38) have supplied the theory of the phenomena for a homogeneous field or where the effects of radiality may be neglected, but recently Archibald (1) has supplied the complete solution for a truly radial or centrifugal field. In both cases it is assumed that diffusion follows Fick's law, completely uninfluenced by the centrifugal field. The movement at any given height in the cell is equal to that due to sedimentation velocity less the diffusion occasioned by the resultant concentration gradients. Hence, initially, diffusion occurs only at the extreme top and bottom of the cell, and sedimentation through any intermediate position is unaffected. This and the following statements are based upon their equations (17, 38, 1) and also the illustrative diagrams for numerical examples.

An appreciable fraction of the material, usually ample for analysis, moves into the lower portion of the cell before the effects of the resulting diffusion gradients reach a level just under the middle of the cell. This is true even where no boundary can form in the upper part of the cell and where by the optical methods only sedimentation equilibrium could be measured. It is likewise true of all cases where a boundary is formed, and hence the sedimentation velocity may be measured from the very beginning, since in *direct* air-driven centrifuges only a fraction of a minute is required to attain full speed. In contrast, with the optical method, the uppermost portion of the cell cannot be used, since time is required for the boundary to be set up, and, as Mason and Weaver have shown, the initial period is occupied in a mere reduction in concentration at the top of the cell before a boundary develops. In both analytical and transparent ultracentrifuges the bottom portion of the cell is not available for measurement of sedimentation velocity, unless it so happens that diffusion is eliminated through the permanent adherence of all sedimenting particles to the periphery of the cell. Summing up, the considerations which limit the applicability of the transparent ultracentrifuge apply in much less degree to the analytical ultracentrifuges, and, if the method of analysis is sufficiently accurate, the latter should be unrestricted in their applicability.

There is one additional diffusional effect to be considered in such opaque ultracentrifuges as those of McBain and Leyda (21), where convection may occur below an immobilized portion of the cell. Diffusion into the immobilized portion must occur. However, it is readily shown by calculation that, since the sedimented material is immediately distributed

throughout a volume of liquid approximately equal to the immobilized portion, this initial diffusion begins at zero rate and only gradually develops into an amount sufficient to affect the measurement of sedimentation velocity. If desired, it can always be taken into account by the usual Fourier calculation.

METHODS OF IMMOBILIZATION

The simplest equipment is the one-piece hollow rotor of Henriot and Huguenard (14), which is directly air-driven. This rotor is photographed in figure 1, and shown in cross section in figure 2. It is unsurpassed for centrifugal force and costs only a few dollars to make. Although admitting

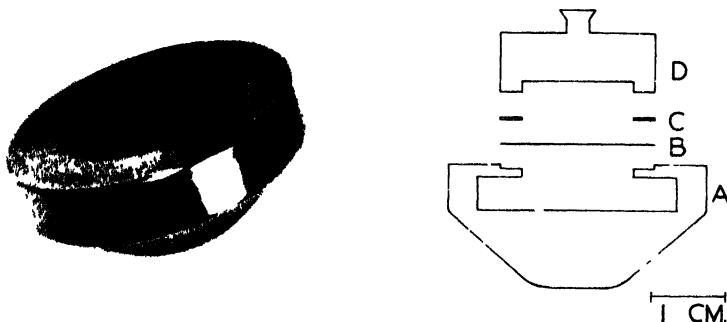


FIG. 1

FIG. 1. One-piece spinning top with depression for window of Cellophane, or of metal, etc

FIG. 2. One-piece hollow steel rotor (A) with cover (B), ring (C), and brass weight (D) to melt on adhesive.

of quantitative results (see later), in this form it is not an ultracentrifuge, but only a convecting centrifuge.

The method of immobilization by a jelly was introduced by McBain and Stuwer (23) and was first applied to measurement of the rate of sedimentation of the jelly structure itself. With 0.3 per cent agar jelly, it gave the same sedimentation rate (65×10^{-13}) as was given (63×10^{-13}) by the transparent ultracentrifuge of McBain and O'Sullivan (22). The swelling pressures of the jelly were also measured. The addition of soap curd was successfully used by McBain and Tostado (20) in measuring the sedimentation equilibrium of sucrose. However, such additions obstruct the rate of sedimentation, as was observed with hemoglobin.

The most general method of immobilization is to place the whole or a

part of the liquid between parallel horizontal disks or washers spaced close enough to immobilize by friction the liquid between the horizontal plates. This method possesses the further advantage that the disks or washers allow ideal unobstructed radial sedimentation over 360°. They must be chosen of material which is strong enough to withstand the centrifugal field and has no action upon the material studied. They have been used for both aqueous and non-aqueous systems.

One might fear a "wall effect" or an influence of the contiguous baffle surfaces. However, according to H. A. Lorenz (16), the force of friction, when a particle with radius r is falling parallel to the wall at a distance l , is increased over that for free fall in an infinite body of liquid taken as unity by the factor

$$1 + \frac{9}{16} \frac{r}{l}$$

Again, according to Faxen (10), the frictional force for a particle falling in the middle between two parallel walls is increased in the ratio

$$\frac{1}{1 - 1.004 \frac{l}{r} + 0.418 \frac{r^3}{l^3} - 0.169 \frac{r^5}{l^5}}$$

Hence, if for egg albumin $r = 22 \text{ \AA}$. and $l = 0.1 \text{ mm.}$, the correction is negligible, affecting only such particles as are within a distance of a few diameters from the wall. This is borne out by the measurements of McBain and Leyda (21) on isoelectric egg albumin, for which they found the value $s_{20} = 3.56 \times 10^{13}$, in agreement with Svedberg's value of 3.55×10^{13} .

While it is always possible to make use of the simple one-piece rotor, it is generally more convenient to place suitable inserts in a two-piece rotor. With the one-piece rotor segments cut radially from washers are piled around the inside of the rotor like brick-work, but with small overlap so that the segments immobilize almost their own volume of liquid between successive horizontal surfaces (19). First, however, a horizontal circular distributing disk is placed in the bottom of the rotor (folding and then flattening it). This permits measurement of rate of sedimentation while the rotor is running by dropping in a heavier liquid, such as carbon tetrachloride, through the distributing disk to force uniformly inwards its own volume of immobilized liquid, which is then removed by a scraping pipet for analysis.

The only convenient and uniformly successful two-piece rotor is that described by McBain and Leyda (21) and shown in figures 3 and 4. These

direct air-driven rotors are distinguished from those of Svedberg and of Beams and Pickels and collaborators by the fact that they maintain themselves at the temperature of the slip-stream of air. The driving air, as in the McBain and O'Sullivan transparent ultracentrifuge, is first passed

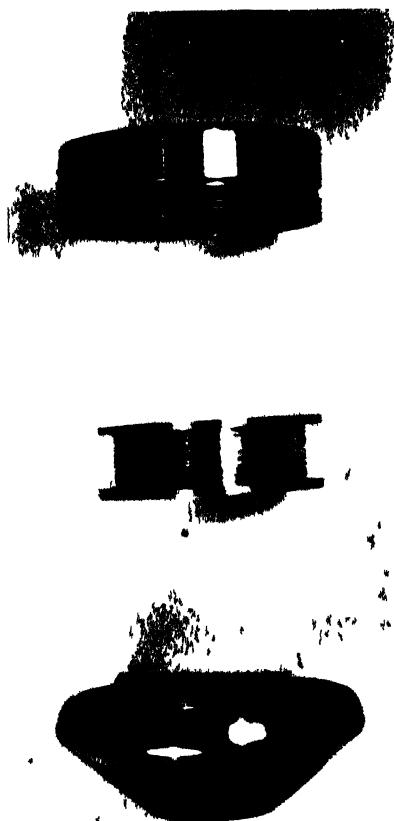


FIG. 3

FIG. 3 Two-piece McBain and Leyda steel rotor with immobilizing insert
FIG. 4 Two-piece McBain and Leyda rotor. A steel cover; B, rotor cone; C, immobilizing disk; D, Phofilm disk; E, metal disk; F, slip ring.

through a copper coil immersed in a thermostat, so that the rotor and its contents are maintained at any desired temperature with the same constancy as the thermostat.

The principle of the McBain-Leyda rotor (figures 3 and 4) is that a

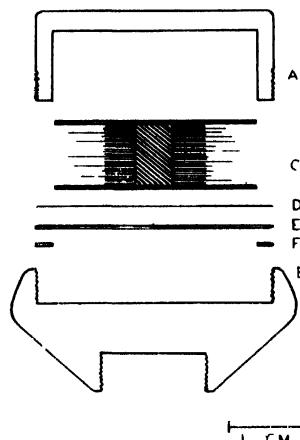


FIG. 4

thrust joint is formed, remaining water- and oil-tight in spite of the centrifugal force subsequently applied. The shell, A, holds the immobilizing insert, C, and the liquid. It screws into the rotor cone, B, its edges forming a seal against a disk of Pliofilm, Cellophane, or other suitable plastic, D, which rests against a rigid loose disk, E, under which is placed a loose metal slip ring, F. The ledge of the rotor cone, with female screw, is made shorter and thicker than the annular walls of the shell, A, the elastic expansion of which into the grooves of the screw assists in maintaining the tight joint during centrifuging. The rotor is assembled upside down and weighed. The liquid is sufficiently immobilized by inserts so that the rotor

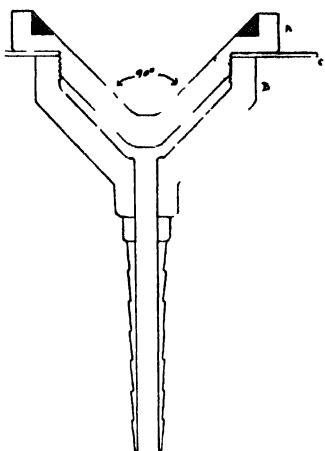


FIG. 5

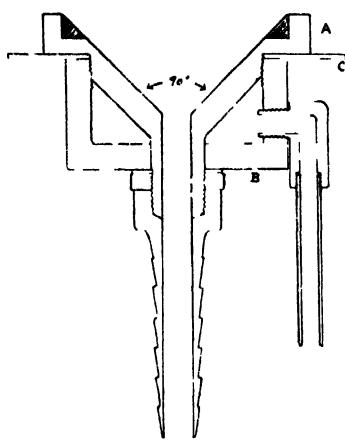


FIG. 6

FIG. 5 Cross section through simple stator A, aluminum cone with brass ring inset and holes for driving air; B, wind-box or manifold; C, aluminum disk to be held by rubber and sponge rubber

FIG. 6. Cross section through stator with central air-inlet to relieve vacuum under rotor to any desired extent.

may be carefully stopped with the fingers and reweighed prior to dismantling for analysis, to make sure that the weight has remained constant to 0.0001 g. With the insert used by Leyda, the total volume of liquid was 2.2 cc., of which the inner half was immobilized.

The typical stators are shown in figures 5 and 6. Figure 6 adopts a suggestion of Beams, namely, that a regulated amount of air be allowed partially to relieve the vacuum under the middle of the rotor. With some rotors this is unimportant, whereas others run more smoothly and up to 10 per cent faster. No motion of the rotor should be visible under the microscope, that is, to 0.001 mm. The stator in figure 5 consists of a

distributing air chamber into which is screwed a cone of 90° angle through whose sides are bored six to twelve holes with a No. 71 drill held at an angle of 65° to the vertical and an angle of 35° to the horizontal. A brass or bronze ring is shown inserted to give longer wear during chance contact, although the rotor is normally carried by the air. The tube leading to

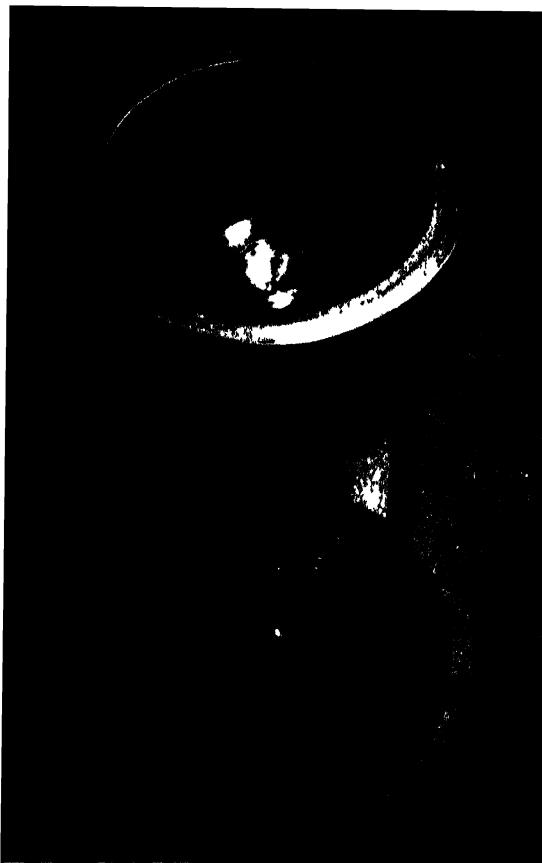


FIG. 7. Stator mounted in lead-lined steel guard (complete)

the stator may be suitably weighted with attached clamps for smoother running. Its upper part should be steadied by pieces of sponge rubber.

A complete assembly is shown in figure 7, with a safety guard for occasions when the rotor flies apart. The original rotor of McBain and O'Sullivan has been in fairly continuous use, running at several thousand R.P.S., during three and a half years without accident.

IMMOBILIZING INSERTS

Various inserts embodying the same principle have been designed for different purposes and have been described elsewhere (19, 21). Here it is sufficient to mention the insert of McBain and Leyda, shown in figures 3 and 4, consisting of a pile of about one hundred circular disks about 0.08 mm. thick and alternately of large and small diameter. The smaller disks thus serve as vertical spacing pieces between the horizontal surfaces of the larger disks where the liquid is immobilized. The whole is held in central position by two large circular plates in the edges of which it is convenient to have cut a couple of notches through which samples of liquid may be withdrawn, with a fine pipet or hypodermic syringe, for analysis. The disks must be made of material indifferent to the solution studied and, to avoid electrical couples, should be made of the same metals throughout if salt is present. The interior of the rotor is best coated with several appli-

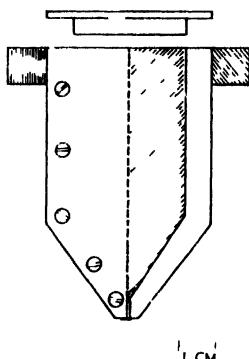


FIG. 8. Metal cuvette for filling and emptying inserts in ordinary centrifuge

cations of Bakelite lacquer, each baked on for 30 min. at 150°C. The disks may be of metal or plastic material, but must withstand the centrifugal force.

Inserts fill spontaneously with liquids of low surface tension, but it may be advisable to place them in a suitable cuvette (figure 8) filled with liquid, for a few moments in an ordinary centrifuge. If the bottom of the cuvette has a slight depression it may also be used for emptying them and collecting the previously immobilized liquid for analysis.

METHOD OF CALCULATION

For sedimentation equilibrium McBain and Tostado (20) have developed a formula suitable for the insert they employed in measuring sucrose (molecular weight: found, 341; theoretical, 342). A corresponding formula has been developed by Leyda for the insert of figures 3 and 4.

For sedimentation velocity by the analytical method, Tiselius, Pedersen, and Svedberg (36) give the formula

$$s_{(\text{observed})} = \frac{1}{2\omega^2 t} \ln \left(1 - \frac{2\Delta}{qxc_0} \right)$$

where ω = angular velocity = 2π (R.P.S.), t is the time in seconds, Δ is the change in amount of substance above or below the level at which the separation is made, x is the distance in centimeters of this level from the center of rotation (radius of larger disks of the McBain and Leyda rotor), q is the cross-sectional area of the cell at this level, and c_0 is the original concentration within the cell. ω , t , x , and q are obtained readily from the dimensions of the insert and observations during the run. In order to apply the analytical method, Tiselius, Pedersen, and Svedberg (36) place a porous partition in the middle of their transparent cell.

McBain and Leyda deduce the formula

$$s_{(\text{observed})} = \frac{2.303(\log x_2 - \log a)}{39.48(\text{R.P.S.})^2 t}$$

where x_2 is the position which the boundary would occupy if no diffusion occurred. This is readily calculated from the dimensions of the insert and from the analysis of the original and final concentrations of the liquid. This possesses for many cases the very great advantage that the absolute concentration need not be known, but only the relative concentrations on any convenient scale.

SEMI-CONVECTIVE METHODS

Elford (6, 8, 9, 35) has used inverted glass, silica, or metal tubes, of 1 to 3 mm. internal diameter, immersed in a commercial or in a Henriot and Hugueneau air-driven centrifuge. Such a design is illustrated in figures 9 and 10, where many parallel holes in one block may be used to give larger volumes. Alternatively we have used with success in this same rotor a capillary tube closed at one end and held in position with a cork, obtaining a boundary with the 70-mm. rotor, where with the same tube inserted in a 37-mm. rotor (figures 3 and 4) a boundary was not formed.

When the tubes are sufficiently narrow and the centrifuge sufficiently free from vibration and thermal gradients, Elford found that his tubes resembled an ultracentrifuge, in that a definite boundary appeared, which could be seen with the naked eye by scattered or fluorescent light and the position of which could be measured after stopping the centrifuge. Fur-

thermore, the boundary moved within the experimental error with the same velocity as that observed in the Svedberg ultracentrifuge.

This is at first surprising for the cell is cylindrical not radial, and some of the sedimentation must be directed obliquely against the walls of the



FIG. 9 Seven centimeter air driven rotor showing recesses for either immersed multiple holes, or simple thick capillary tube closed at one end with large cover for streamlining

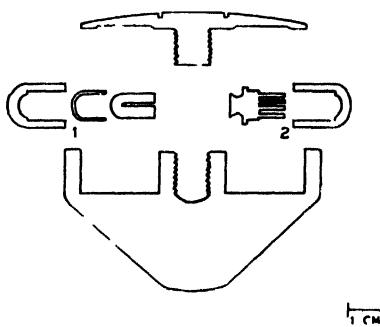


FIG. 10 Cross section of 7 cm air-driven rotor showing recesses for (1) simple thick capillary tube closed at one end or (2) immersed multiple holes

tube. Thus a sheath of liquid in contact with the walls of the tube must move steadily downwards and correspondingly elevate the central contents, including the boundary. Our explanation of the fact that correct

results are obtained is that the movement is slow and controlled and confined to a layer next to the walls, which is only a few molecules or particles thick and is therefore of a volume negligible in comparison with the contents of the tube.

McBain (18) pointed out that, since the sedimentation velocity depends upon the actual specific volume of the particle as well as the density of the solution, the actual density of the particle may be determined by suitable additions to change the density of the solution. Giving a formula confirmed by Kraemer and Lansing (15), he showed that sedimentation velocity is as little affected by hydration in a binary system, consisting of solvent and only one solute, as is any osmotic property such as the lowering of the freezing point. It is far different when further additions, such as buffers, are present, for the particle may be made to sediment either upwards or downwards, and thus its actual density is determined. From the actual, as compared with the partial, specific volume of the anhydrous material the actual hydration or solvation may be found. Using this principle Elford (7) has found the density of vaccinia virus to be 1.17 ± 0.02 and that of two of the largest bacteriophages—*Staphylococcus* "K" and *Coli* "WLL" to be 1.22 ± 0.02 . Many other similar determinations of density have now been made, such as those made by McIntosh and Selbie (24, 25).

Schlesinger (30, 31) has used a successful device for converting a Sharples supercentrifuge into a semi-convectionless ultracentrifuge. Five cubic centimeters of virus solution gelatinized with dilute agar lines the closed bowl to a depth of 0.18 mm. Another 5 cc. is then added. The film is so thin that convection does not occur, thus allowing both rate and equilibrium to be measured. Virus of foot-and-mouth disease is measured after 3 min. An antibody required only 30 min. for sedimentation equilibrium. It is necessarily assumed that the agar jelly is of such concentration that it neither swells nor sediments. This, however, can be verified by direct experiment and by adjusting the concentration of agar to the requisite value. Any influence of the agar on the absolute rate has to be tested by comparison in some other ultracentrifuge. Sedimentation equilibrium is of course unaffected.

THE CONVECTIVE PROCEDURE

The Bechhold-Schlesinger convective procedure (2, 29) was originated in 1931, and during the last eight years (5, 8, 9, 11, 12, 13, 24, 25, 27, 28, 35, 37) various qualitative and semi-quantitative observations of its occurrence have been made in the author's laboratory at Stanford (as, for example, with methylene blue, etc.), in that of Beams at Virginia, and also by Gratia in Belgium, using the simplest form of the one-piece hollow rotor of Henriot and

Huguenard (14). If the outer periphery of a centrifuge is of such a nature that through porosity, adhesion, or coagulation every particle sedimented against that periphery remains therein, the liquid remains of uniform composition which, however, decreases asymptotically with time in accordance with a formula of Bechhold and Schlesinger. From this one may calculate sedimentation velocity. It has not been sufficiently emphasized in recent years that, when the centrifugation is carried to near exhaustion of the solution, the final concentration becomes extraordinarily sensitive to particle size. This may sometimes be by far the most accurate method of examination.

It is interesting that if the rotor is filled with liquid to the center and the outer periphery has the property of holding all sedimented particles, there is no difference between a convectionless ultracentrifuge and a thoroughly stirred centrifuge.

The convective method has been used by McIntosh and Selbie (24, 25) for quantitatively determining the sedimentation velocity and actual densities of bacteria, viruses, bacteriophages, and oxyhemoglobin. For example, they obtained a diameter of 56 Å. for oxyhemoglobin, identical with that quoted from Svedberg. Their equipment was very similar to that in figures 9 and 10. The density of spores of *B. subtilis* was found to be 1.46, and that of phages and *Staphylococcus aureus* 1.25. In solutions of higher density they sedimented upwards.

Paic (28, 5, 27, 37) has suggested the standardization of the sedimentation velocity constant of an unknown by comparison with that of a known in the same centrifuge, and has deduced the theoretical ratios with and without convection, respectively. However, he found the relation

$$s/s' = \varphi \cdot t'/t$$

where t and t' are the respective times required to reduce the concentration of the known and the unknown by the same fraction in the same centrifugal field, and φ is a simple allowance for the respective partial specific volumes.

Gratia (11, 12, 13) and Paic (28, 5, 27, 37) and their respective collaborators have studied serums, venoms, antibodies, antitoxins, antigens, lysins, and syphilitic reagents.

REFERENCES

- (1) ARCHIBALD: Phys. Rev. **53**, 746 (1938); **54**, 371 (1938).
- (2) BECHHOLD, H., AND SCHLESINGER, M.: Biochem. Z. **236**, 392 (1931); Z. Hyg. Infektionskrankh. **112**, 668 (1931); **115**, 342, 354 (1933); Phytopath. Z. **6**, 627 (1933).
- (3) BURK, N. F.: J. Biol. Chem. **121**, 373 (1937); **120**, 63 (1937).
- (4) BURK, N. F., AND GREENBURG, D. M.: J. Biol. Chem. **87**, 197 (1930).

- (5) DEUTSCH, V., AND LOMINSKI, I.: Bull. acad. méd. **118**, 220 (1937).
- (6) ELFORD, W. J.: Brit. J. Exptl. Path. **17**, 399 (1936).
- (7) ELFORD, W. J.: Proc. Roy. Soc. (London) **B125**, 309-10 (1938).
- (8) ELFORD, W. J., AND ANDREWES, C. H.: Brit. J. Exptl. Path. **17**, 422 (1936).
- (9) ELFORD, W. J., AND GALLOWAY, I. A.: Brit. J. Exptl. Path. **18**, 155 (1937).
- (10) FAXEN, H.: Ann. Physik [4] **68**, 89 (1922).
- (11) GRATIA, A.: Bull. soc. chim. biol. V^e congrès **18**, 208 (1936); Compt. rend. soc. biol. **117**, 1228 (1934).
- (12) GRATIA, A., AND GORECSKI: Compt. rend. soc. biol. **126**, 900, 1252 (1937).
- (13) GRATIA, A., AND NELIS, P.: Compt. rend. soc. biol. **127**, 840 (1938).
- (14) HENRIOT AND HUGUENARD: Compt. rend. **180**, 1389 (1925); J. phys. radium **8**, 433 (1927).
- (15) KRAEMER, E. O., AND LANSING, W. D.: J. Am. Chem. Soc. **55**, 4319 (1933); J. Phys. Chem. **39**, 165 (1935).
- (16) LORENZ, H. A.: Abhandl. theor. Physik **1**, 23 (1906).
- (17) MASON, M., AND WEAVER, W.: Phys. Rev. **28**, 412 (1924).
- (18) MCBAIN, J. W.: J. Am. Chem. Soc. **55**, 4319 (1933).
- (19) MCBAIN, J. W.: J. Phys. Chem. **42**, 1063 (1938), figure 3.
- (20) MCBAIN, J. W., AND ALVAREZ-TOSTADO, C.: Nature **139**, 1066 (1937); J. Am. Chem. Soc. **60**, 2489 (1937).
- (21) MCBAIN, J. W., AND LEYDA, F. A.: J. Am. Chem. Soc. **60**, 2998 (1938); Nature **141**, 913 (1938).
- (22) MCBAIN, J. W., AND O'SULLIVAN, C.: J. Am. Chem. Soc. **57**, 780, 2631 (1935).
- (23) MCBAIN, J. W., AND STUEWER, R. F.: Kolloid-Z. **74**, 10 (1936).
- (24) MCINTOSH: J. Path. Bact. **41**, 215 (1935).
- (25) MCINTOSH AND SELBIE, F. R.: Brit. J. Exptl. Path. **18**, 162 (1937).
- (26) MEYER, K. H., AND LUDHEMANN, R.: Helv. Chim. Acta **18**, 307 (1935).
- (27) PAIC, M.: Compt. rend. **207**, 629, 1074 (1938).
- (28) PAIC, M., AND CHOROKHOFF, M.: Compt. rend. soc. biol. **126**, 877 (1937).
- (29) SCHLESINGER, M.: Z. Hyg. Infektionskrankh. **114**, 161 (1932); Biochem. Z. **284**, 6 (1933); Kolloid-Z. **67**, 135 (1934); Biodynamica **1935**, 1.
- (30) SCHLESINGER, M.: Nature **138**, 549 (1936).
- (31) SCHLESINGER, M., AND GALLOWAY, I. A.: J. Hyg. **37**, 445, 463 (1937).
- (32) STEINHARDT, J.: J. Biol. Chem. **123**, 543 (1938).
- (33) SVEDBERG, T.: Ind. Eng. Chem., Anal. Ed. **10**, 116 (1938).
- (34) SVEDBERG, T.: Ind. Eng. Chem., Anal. Ed. **10**, 118 (1938).
- (35) TANG, F. F., ELFORD, W. J., AND GALLOWAY, I. A.: Brit. J. Exptl. Path. **18**, 269 (1937).
- (36) TISELIUS, A., PEDERSEN, K. O., AND SVEDBERG, T.: Nature **140**, 848 (1937).
- (37) VIUCHANGE, J., AND PAIC, M.: Compt. rend. **206**, 1318 (1938).
- (38) WEAVER, W.: Phys. Rev. **27**, 499 (1926); Z. physik **43**, 296 (1927); **49**, 311 (1928).
- (39) WEBER, H. H., AND STÖVER, R.: Biochem. Z. **259**, 269 (1933).
- (40) WILLIAMS, J. W., AND WATSON, C. C.: Nature **139**, 506 (1937).

THE PHYSICAL CHEMISTRY OF TOBACCO MOSAIC VIRUS PROTEIN¹

MAX A. LAUFFER AND W. M. STANLEY

The Department of Animal and Plant Pathology of The Rockefeller Institute for Medical Research, Princeton, New Jersey

Received January 30, 1939

INTRODUCTION

Tobacco mosaic, a disease which manifests itself by causing a mottling and a distortion of the leaves of tobacco plants, was first recognized over eighty years ago. In 1892 Iwanowski (19) observed that the juice of plants diseased with tobacco mosaic remained infectious after passing through a Chamberland filter, a device capable of holding back all of the living organisms then known. After repeating and confirming Iwanowski's experiments, Beijerinck (5), in 1898, realized that the infectious agent differed from ordinary bacteria and described it as being a "*contagium vivum fluidum*." Tobacco mosaic was thereby established as the first disease recognized as being caused by an agent now known as a virus.

In 1935, by a process consisting of fractionation by chemical means, a crystalline protein was isolated from the juice of Turkish tobacco plants infected with tobacco mosaic virus (40). This protein has been shown to be in chemical combination with a nucleic acid closely resembling yeast nucleic acid and is, therefore, a nucleoprotein (2, 27). It has definite and specific chemical composition and physical characteristics. It gives the usual protein color reactions, is precipitated by the usual protein precipitating agents, has characteristic heat and pH stability ranges, and is denatured under certain definite conditions. Very dilute solutions give a specific precipitin reaction with antiserum to the protein. The protein is insoluble at its isoelectric point and in 20 per cent ammonium sulfate solutions. It crystallizes in definite needle-shaped crystals in the paracrystalline or mesomorphic state, visible only with the microscope.

All of the evidence available at present indicates that this crystallizable nucleoprotein is the active disease-causing agent. The virus protein

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

isolated from many different batches of diseased Turkish tobacco plants and from other plant species infected with the virus possesses the same chemical, physical, biological, and immunological properties. This protein is capable of infecting other plants, will multiply in them, and can be isolated from such diseased plants. There is no reason to believe that the protein is not pure, for its chemical, physical, biological, and immunological properties remain unchanged following fractionation of the protein by various procedures. The ultracentrifugally isolated preparations of the material are completely homogeneous with respect to sedimentation constant and electrochemical behavior (14, 53). It was found impossible to demonstrate the presence of an impurity in such preparations even by the sensitive precipitin and anaphylactic tests. It has never been found possible to separate the virus activity from the protein by any one of several procedures. The ultraviolet light absorption spectrum of the protein agrees essentially with the destruction spectrum of virus activity (2, 10, 26). The pH stability range of the protein was found to coincide with that of virus activity. Partial or complete denaturation of a preparation by each of several procedures was found to result in a corresponding loss of virus activity. Finally, it is possible not only to inactivate and reactivate the virus protein, but also to demonstrate that the inactivation and reactivation are accompanied by simultaneous changes in the structure of the protein molecule (36). For complete and detailed descriptions of the properties of the tobacco mosaic virus protein, the reader is referred to other reviews (41, 42, 43, 44, 45). In this paper a summary of many of the physical and physicochemical properties of the protein is presented, and an attempt is made to unify the interpretation of these properties.

THE SHAPE AND SIZE OF THE TOBACCO MOSAIC VIRUS PROTEIN PARTICLES

Optical studies

Takahashi and Rawlins (48) first showed that the juice from tobacco plants diseased with tobacco mosaic exhibited stream double refraction, and it has since been demonstrated that solutions of the purified tobacco mosaic virus protein show the phenomenon to a marked degree (2, 3, 25, 49) (figure 1). This property may be the result of one or more of the following three factors: (1) the photoelastic effect, (2) the orientation of plate-like bodies in the flowing stream, and (3) the orientation of rod-like particles. The possibility that stream double refraction may result from the first (25) or second of these factors (48) has been eliminated; hence it may be concluded that the double refraction of flow shown by the tobacco mosaic virus protein is due to the orientation of rod-shaped particles in the flowing stream.

Wiener (52) has shown on theoretical grounds that it is possible to obtain

double refraction by orienting perfectly isotropic rod-shaped particles parallel to each other in a liquid medium, if the rods have a different refractive index from that of the medium. This is called morphic double refraction. If the rod-shaped particles have intrinsic double refraction, the double refraction of the system of oriented rods in a liquid medium is the sum of the morphic and intrinsic double refractions. By dispersing the rods in a medium having the same refractive index as the rods, the morphic double refraction can be completely eliminated, and any remaining double refraction is the intrinsic double refraction of the rods. It has been shown by the method indicated that the tobacco mosaic virus protein particles

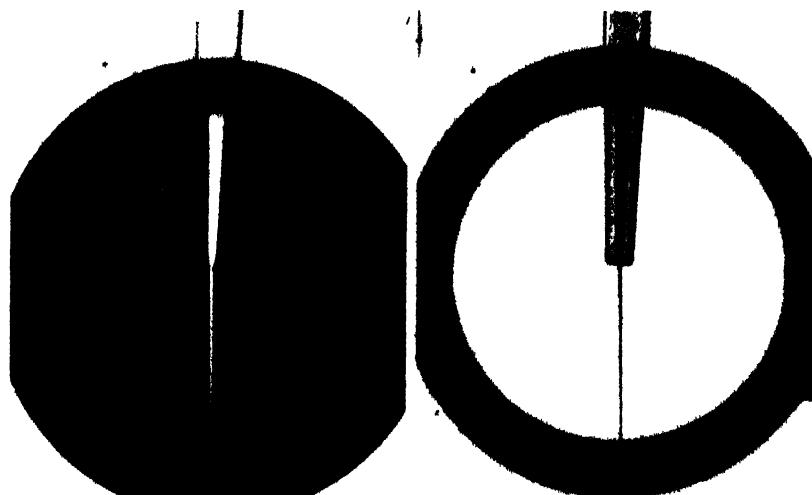


FIG. 1 Left: doubly refracting stream of tobacco mosaic virus protein solution photographed between crossed Polaroid plates arranged so that the vibration direction of each plate makes an angle of 45° with the direction of flow. Right: same system photographed between parallel Polaroid plates (From Lauffer and Stanley (25)).

possess very little if any intrinsic double refraction and that the double refraction shown by solutions of the oriented molecules must be largely morphic double refraction (figure 2) (22). This result is entirely consistent with the observation of Bawden *et al.* (3) that dried films of the protein show very much less double refraction than wet films.

Solutions of the protein also show electrical birefringence (3). The birefringence is positive with respect to the direction of the electrical field, showing that the particles orient themselves parallel to the direction of the field (24).

Bawden and Pirie (2) first noticed that, upon standing, some solutions

of virus protein separate into two layers, of which the lower more concentrated one is liquid crystalline (figure 3). The jelly-like pellets obtained by ultracentrifuging solutions of the protein have also been shown to be

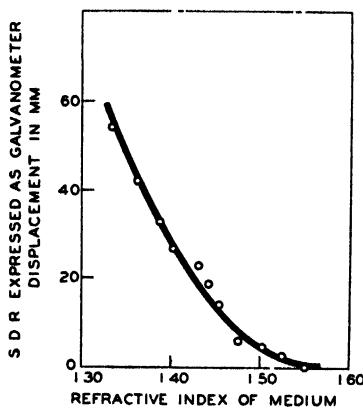


FIG. 2. The dependence of stream double refraction of tobacco mosaic virus protein upon the refractive index of the solvent. (From Lauffer (22)).

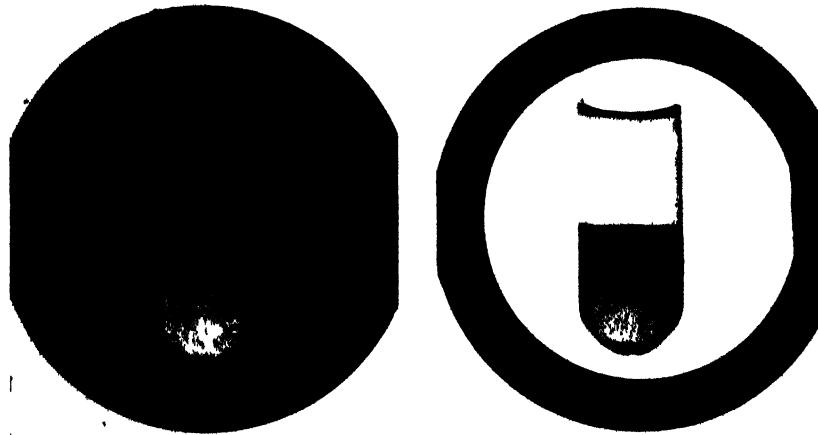


FIG. 3. Left: sample of tobacco mosaic virus protein solution which has separated into two layers. The photograph, taken with the aid of crossed Polaroid plates, shows that the bottom layer material is spontaneously doubly refracting or liquid crystalline. Right: same system photographed between parallel Polaroid plates. (From Lauffer and Stanley (25)).

typically liquid crystalline (22). This property of liquid crystallinity is generally associated with materials having rod-shaped molecules, and may therefore be regarded as constituting additional evidence of the rod-

like character of the tobacco mosaic virus protein particles. The layering behavior illustrated in figure 3 is entirely analogous to that described previously by Zocher and Jacobsohn (57) for vanadium pentoxide sols.

As one should expect of a substance having rod-shaped particles, solutions of the tobacco mosaic virus protein show the Ganz depolarization effect (22). The depolarization of the scattered light observed in the case of this protein is of the type regarded as being due to asymmetry or anisotropy of particles of the disperse phase, rather than to particle size being beyond a critical limit. All of these pieces of evidence, then, in addition to the fact that the virus protein crystallizes in needle-shaped mesomorphic crystals, leave little doubt but that the molecules of the tobacco mosaic virus protein are rod-shaped bodies.

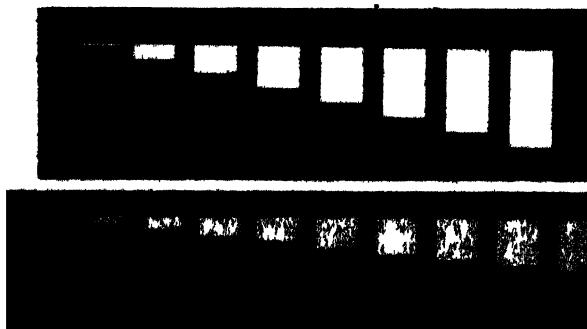


FIG 4 Upper: sedimentation photograph of ultracentrifugally isolated tobacco mosaic virus protein showing single sharp boundary indicative of homogeneity. Lower: sedimentation photograph of an ultracentrifugally isolated sample which had stood in contact with electrolytes, showing two sharp boundaries indicating the presence of two molecular species (Courtesy of Wyckoff)

Ultracentrifugation studies

The first sedimentation studies on the tobacco mosaic virus were made on infectious juice by Bechhold and Schlesinger (4) before the virus was isolated as a purified protein. They followed the sedimentation by means of measurement of virus activity and obtained data from which a sedimentation constant may be calculated that is in good agreement with that found later for the purified protein. Sedimentation studies on the purified tobacco mosaic virus protein have been made by Eriksson-Quensel and Svedberg (14) and by Wyckoff (53, 54, 56). It was found that the protein isolated and purified by differential centrifugation gives a single sharp boundary in the ultracentrifuge (53) (figure 4), with a sedimentation constant, S_{20} , of 174×10^{-12} .

It is not possible to interpret directly this sedimentation constant in

terms of molecular weight, because it is known that the molecules of the tobacco mosaic virus protein are not symmetrical. In order to calculate the molecular weight from these studies, it is necessary to know the dissymmetry factor of the protein. This is usually obtained from sedimentation equilibrium measurements (46), but, because of technical difficulties, it was not found possible to obtain satisfactory results by this method (14).

Viscosity studies

It is possible to obtain an idea of the dissymmetry of rod-shaped particles from studies of the viscosity of solutions or suspensions of these particles. Viscosity studies on this protein have been reported by Stanley (43), by Frampton and Neurath (15), and by Lauffer (21, 23). In figure 5 the data of Frampton and Neurath and of Lauffer are presented graphically. It is seen that viscosity is a linear function of concentration for very dilute

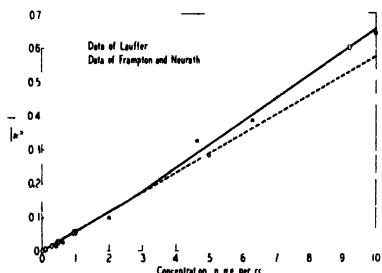


FIG. 5. The relationship between specific viscosity and concentration of tobacco mosaic virus protein. (Drawn from data of Frampton and Neurath (15) and of Lauffer (21, 23)).

solutions of the virus protein, but that this linearity does not hold for solutions as concentrated as 1 per cent. The results of the viscosity studies on the virus protein in urea solutions recently reported by Frampton and Saum (15a) have no direct relationship to the present problem, for they refer to degradation products which have been found not to possess virus activity (24).

Kuhn (20) has derived the following equation expressing the viscosity of a solution or suspension of rod-shaped particles as a function of the relative volume and the relative dimensions of the particles of the disperse phase:

$$\frac{\eta}{\eta_0} = 1 + 2.5G + \frac{G}{16} \left(\frac{b}{a} \right)^2 \quad (1)$$

This equation was derived for a rod-like model composed of rigidly joined spheres. Guth (16) has derived the same expression for a model consisting

of an elongated ellipsoid of revolution. G is the volume of the dispersed material per cubic centimeter of solution, η/η_0 is the relative viscosity of solution, and b/a is the ratio of length to diameter of the rods of the disperse phase. An alternate equation has been derived by Eisenschitz (12) for elongated ellipsoids of revolution. When $b \gg a$, this equation takes the form,

$$\frac{\eta}{\eta_0} = 1 + G \frac{\left(\frac{b}{a}\right)^2}{15 \left(\ln 2 \frac{b}{a} - \frac{3}{2} \right)} \quad (2)$$

The Eisenschitz equation, in its expanded form, reduces to the Einstein viscosity equation for $(b/a) = 1$, just as does the Kuhn-Guth equation. Using equation 1, assuming little or no hydration of the protein, and taking 0.73 as the specific volume of the protein (2, 43), the value of b/a calculated from the limiting slope of the viscosity-concentration curve is 35. Using equation 2 and the viscosity data presented in figure 5, one obtains a value of 63 for b/a .

An equation describing the relationship between the ratio of minor to major axes of an ellipsoid of revolution and the dissymmetry factor of Svedberg, (D_0/D) , has been presented by Herzog, Illig, and Kudar (18) and Perrin (34). It may be used in the following form when $b \gg a$:

$$(D/D_0) = \sqrt{\frac{\left(\frac{a}{b}\right)^{2/3}}{1 - \left(\frac{a}{b}\right)^2} \ln \frac{1 + \sqrt{1 - \left(\frac{a}{b}\right)^2}}{\frac{a}{b}}} \quad (3)$$

The diffusion constant of particles of any shape, D , is related to the sedimentation constant, S , according to the following equation:

$$D = \frac{RTS}{M(1 - Vd)} \quad (4)$$

where R is the gas constant, T is the absolute temperature, M is the molecular weight, V is the partial specific volume of the dispersed particles, and d is the density of the solvent (46). The diffusion constant of spherical particles, D_0 , is given according to the Einstein equation:

$$D_0 = \frac{RT}{6\pi\eta_0 N} \left(\frac{3MV}{4\pi N} \right)^{1/3} \quad (5)$$

where N is Avogadro's constant and M is the molecular weight. Since D_0/D may be evaluated by equation 3, the molecular weight can be calcu-

lated from sedimentation data by combining equations 4 and 5. A value of 42.6×10^6 is obtained for the molecular weight of the tobacco mosaic virus protein by this method when the viscosity data are interpreted according to equation 1. This would correspond to rod-shaped particles 12.3 m μ in diameter and 430 m μ in length. If the value b/a obtained from the viscosity data employing equation 2 is used in conjunction with equations 3, 4, and 5, a value of 63.2×10^6 is obtained for M , corresponding to particles 11.5 m μ in diameter and 725 m μ in length.

Diffusion

Another means of evaluating the molecular weight of material having asymmetrical molecules is by combining diffusion and sedimentation measurements in the manner indicated by equation 4 (46). Diffusion studies on purified solutions of tobacco mosaic virus have been made by Waugh and Vinson (51) and by Hills and Vinson (18a). It is probable that the diffusion constants reported by the latter workers are too high because of failure to suppress the accelerating effect of smaller oppositely charged ions. Their virus solutions, which contained small amounts of electrolytes, were allowed to diffuse through a porous membrane into distilled water. Neurath and his associates (15, 32, 33) have measured the diffusion coefficient of chemically isolated virus protein, using the refractory method of Lamm. They found a value of about 3×10^{-8} for D_{25° . From equation 4 it follows that this value corresponds to a particle with a molecular weight of 59.6×10^6 . By the use of equations 3 and 5, it can be shown that such particles should have a ratio of length to width of 58, corresponding to dimensions of 675 m μ in length and 11.6 m μ in diameter (32, 33).

It is readily apparent that the molecular weight may be estimated from viscosity and diffusion data by the use of equations 1 or 2, 3, and 5 (32, [33, 35]). If the data are treated in the indicated manner using equation 1, a value for the molecular weight of 116.5×10^6 is obtained, corresponding to particles 600 m μ long and 17.2 m μ in diameter. However, if equation 2 is used instead of equation 1, the calculated molecular weight is found to be 53.4×10^6 , corresponding to particles 690 m μ long and 10.9 m μ in diameter.

The rotational diffusion coefficient

Boeder (9) has derived an equation in which the average orientation of rod-shaped particles in a flowing liquid is expressed as a function of the ratio of the velocity gradient tending to align the particles to the rotational diffusion constant. The latter is a measure of the tendency of the particles to become randomly oriented. Kuhn (20) has expressed the relationship

between the rotational diffusion constant, θ , and the length of a rod-shaped particle, b , by the equation:

$$\theta = \frac{8kT}{\pi\eta b^3} \quad (6)$$

where k is the Boltzmann constant. From quantitative measurements of double refraction of flow and the average orientation of the particles, Mehl (31), following the method of Boeder, estimated the rotational diffusion constant for chemically isolated tobacco mosaic virus protein to be 25 at 2°C. By substituting this value into equation 6, he obtained the value of about 610 m μ for the length of the particle, a value which is of the same order of magnitude as those obtained from the other data considered.

Filtration of the tobacco mosaic virus protein

The whole body of knowledge of viruses began with the observation that the infectious agent of tobacco mosaic would pass through a Chamberland filter (5, 19). In 1916 Allard (1) reported that, although the virus of tobacco mosaic would easily pass a Berkefeld filter, it was held back by a Livingstone atmometer porous cup. Later Duggar and Karrer (11) reported that the virus had about the same filterability as colloidal hemoglobin, and suggested a value of 30 m μ for the diameter of the particles of the infectious agent. MacClement and Smith (30) found that two strains of tobacco mosaic would pass collodion filters of pore diameter of 51 m μ , while a third strain, aucuba mosaic, would pass through pores 112 m μ in diameter but would be held by those 100 m μ in diameter. On the basis of Elford's studies (13), these results correspond to particle diameters of 17 to 25 m μ and 37 to 56 m μ , respectively. Thornberry (50) found that eight strains of tobacco mosaic virus, including the aucuba mosaic virus, as well as several other plant viruses, when diluted with a solvent containing nutrient broth, filtered through pores as small as 45 m μ , indicating that they all have diameters of about 15 to 22 m μ . All of the studies thus far reported were on plant juice containing the virus. Thornberry also studied tobacco mosaic virus protein purified by the modified lead acetate method of Stanley and found that this material would go through pores 33.8 m μ in diameter. In contrast with this observation, Bawden and Pirie (2) and Smith and MacClement (39) have reported that purified tobacco mosaic virus protein will not pass through a collodion membrane with a pore size as great as 450 m μ . They regard this as evidence of an aggregation of the protein due to purification. However, results obtained with protein isolated by differential centrifugation in the Princeton laboratory of The Rockefeller Institute for Medical Research indicate that this purified protein will pass freely through membranes with a pore size of 450 m μ , al-

though a diminished rate of passage in some membranes of smaller pore size has been noticed (29, 45). Smith and MacClement reported that in an electrical field the tobacco mosaic virus will pass through a membrane having an average pore diameter of $13 \text{ m}\mu$ (39).

If it is assumed that rod-shaped particles pass through the pores of the filter lengthwise, and that the ratio of pore diameter to particle diameter obtained by Elford (13) for symmetrical particles holds for the diameter of the rods, the results of the filtration of the virus protein would indicate a diameter of about $15 \text{ m}\mu$,—a value in good agreement with that found by other means.

X-ray studies

X-ray diffraction patterns of the tobacco mosaic virus protein have been studied in the laboratories of Bernal (3, 6, 7) and of Wyckoff (55). Bernal and Fankuchen reported studies on the virus protein dissolved in water, on the liquid crystalline "bottom layer" material (see figure 3), on films of the protein obtained by evaporation and designated as wet and dry gels, respectively, and on oriented crystals of the protein. Wyckoff and Corey studied the diffraction patterns of crystals and of pellets obtained with the ultracentrifuge. The findings from the two laboratories agree on the intramolecular structure of the material. Bernal and Fankuchen regard the unit, which, because of its physical properties, is regarded as a molecule, as having an *internal* crystalline structure of great regularity, analogous to the structure of crystalline proteins. The lines that indicate this internal structure are found in the virus in all of the states studied, including solutions, and therefore must define the molecule rather than the crystal.

A lateral spacing of $15.2 \text{ m}\mu$ is found in the "dry gel" studied by Bernal and Fankuchen. The wet gel gives corresponding spacings varying around $21 \text{ m}\mu$ depending upon the composition, the liquid crystalline bottom layer shows analogous spacings of between 30 and $47 \text{ m}\mu$, again depending on the concentration of the protein, and the ordinary solutions show no corresponding spacings. The process of drying a solution of the virus protein, then, consists of the removal of water until the solution reaches the concentration where the particles line up parallel to each other, giving a liquid crystal with rather wide intermolecular spacings. Further evaporation results in the removal of water from between these particles, gradually decreasing the intermolecular spacings until the dry gel state is reached, when most of the water has been removed. The intermolecular distance then becomes $15.2 \text{ m}\mu$. Since this spacing is also obtained in the dried crystals, it probably represents the closest possible spacing of the molecules, and therefore provides an estimate of the molecular diameter.

Whereas Bernal and Fankuchen found this lateral spacing of $15.2 \text{ m}\mu$ in

the oriented crystals, they were unable to find evidence of intermolecular regularity of less than $120 \text{ m}\mu$ in directions other than at right angles to the length of the crystal. They accordingly suggested that the crystals are made up of rod-shaped molecules arranged parallel in hexagonal symmetry with respect to cross section, the intermolecular distances being $15.2 \text{ m}\mu$, and without intermolecular regularity in the direction of the long axis. In figure 6 this picture is presented diagrammatically.

It should be emphasized that the x-ray diffraction data indicate a particle $15.2 \text{ m}\mu$ in diameter and probably more than $120 \text{ m}\mu$ in length. These figures for the dimensions of the particle are in good agreement with those arrived at from viscosity and sedimentation studies.



FIG. 6. The arrangement of the rod-shaped molecules of tobacco mosaic virus protein in the needle-shaped crystals. Left: longitudinal diagram showing parallel orientation of particles and no intermolecular regularity in lengthwise direction. Right: cross-sectional diagram showing hexagonal intermolecular symmetry. (Drawn from Bernal and Fankuchen (7)).

Adsorption on solids

Langmuir and Schaefer (20a) adsorbed tobacco mosaic virus protein on an A-layer of stearic acid conditioned with aluminum chloride and found by means of interference colors that the film was about $30 \text{ m}\mu$ thick. They found further that a surface conditioned by a monolayer of egg albumin takes up an adsorbed film of the virus protein having a thickness of $12.5 \text{ m}\mu$. This latter they suggested, may be the thickness of the molecules lying flat on the surface, and indeed this value is in agreement with those obtained by the other methods just discussed.

General considerations

In summarizing the knowledge of the size of the tobacco mosaic virus protein molecule, it becomes evident that several experimentally determined constants for the protein, as well as adsorption, ultrafiltration, and x-ray diffraction data, are available. The sedimentation constant at $20^\circ\text{C}.$, S_{20} , is 174×10^{-13} , the specific viscosity per gram per milliliter is 57.9, the diffusion constant at $25^\circ\text{C}.$, D_{25} , is 3×10^{-8} , the rotational

served for the purified virus protein, regard it as being an aggregated form of the virus as it occurs in the natural state and in untreated plant juice. This seems a reasonable interpretation of their data. However, the rather meager sedimentation data at present available indicate that this interpretation is not valid for the homogeneous ultracentrifugally isolated material. The sedimentation constant of the virus protein in a sample of untreated plant juice was found (56) to be the same as that of the protein later isolated from the same juice. If it is to be supposed that an aggregation took place in the purification process, the virus particles in the juice must have been at most half as large as the particles in purified preparations. The type of aggregation which would affect the sedimentation constant least is an end to end association of the hypothetical sub-units.

TABLE 2

Sedimentation constants of molecular species in tobacco mosaic virus proteins prepared at various intervals after plant inoculation
(Data of Wyckoff (53))

| TIME AFTER INOCULATION weeks | SEDIMENTATION CONSTANTS ($S \times 10^{12}$) | |
|---------------------------------|------------------------------------------------|------------------------------|
| | Ultracentrifuged samples | Chemically extracted samples |
| 1 | 174.3 | 177.1 |
| 2 | 180.0 | 179.3, 203.5 |
| 3 | 177.6 | 170.6, 198.7 |
| 4 | 177.1 | 172.1, 201.9 |
| 5 | 173.1, 198.8 | 172.7, 201.1 |
| 6 | 172.9, 199.7 | 170.9, 195.4 |
| 7 | 175.2, 198.9 | 175.2, 196.6 |
| 8 | 175.3, 197.9 | 178.3, 203.2 |
| 13 | 174.0, 205.3 | |

By the same method as was just used, it was shown that sub-units which could associate to give whole molecules of molecular weight 42.6×10^6 , ratio of length to diameter of 35, and S_{20° of 174×10^{-12} would have a molecular weight of 21.3×10^6 , a ratio of length to diameter of 17.5, and a sedimentation constant of 145×10^{-12} (23). Such a value differs sufficiently from 174×10^{-12} so that it could easily be differentiated. Any other set of assumptions would lead to an even lower sedimentation constant for the virus in juice. Since no more slowly moving boundary was found in the sample of juice studied, it must be concluded that the virus molecules in this sample of plant juice were of the same weight and size as those in purified samples prepared by careful differential centrifugation.

FURTHER PHYSICAL-CHEMICAL STUDIES

Absorption spectrum

The ultraviolet light absorption spectrum of crystalline tobacco mosaic virus protein has been studied by Lavin and Stanley (26) and by Bawden and Pirie (2). It shows a spectrum similar to that of other proteins, excepting that it has a maximum at 2650 Å. instead of at 2800 Å. This is an important observation, inasmuch as the inactivation of the virus by irradiation is at a maximum for wave length approximating 2650 Å. (10).

Surface spreading

Using the Langmuir trough, Seastone (37, 38) studied the surface spreading of tobacco mosaic virus protein. He found that the protein would not spread on dilute solutions of electrolytes (0.02 M) except in the

TABLE 3

The spreading of tobacco mosaic virus protein and of egg albumin on the surfaces of ammonium sulfate solutions
(Data of Seastone (38))

| CONCENTRATION OF AMMONIUM SULFATE (PER CENT SATURATION) | FILM AREA AS M^2 PER MILLIGRAM OF PROTEIN | |
|------------------------------------------------------------|----------------------------------------------------|-------------------------------------|
| | Tobacco mosaic <i>square meters</i> | Egg albumin <i>square meters</i> |
| 0 | 0.000 | 0.1 |
| 12.5 | 0.008 | 0.8 |
| 25.0 | 0.024 | 1.0 |
| 50.0 | 0.020 | 1.0 |
| 75.0 | 0.025 | 1.0 |
| 90.0 | 0.022 | 1.0 |

region of pH 1, where it spread to a small extent. His results show, however, that the protein will spread on concentrated ammonium sulfate solutions to a maximum extent of from 0.020 to 0.025 square meter per milligram (38). This is in contrast to the square meter per milligram occupied by films of such proteins as egg albumin and pepsin made in a similar manner (17). Evidently the tobacco mosaic virus protein does not unfold into a monolayer of a thickness corresponding to that of one amino acid, as do the proteins of lower molecular weight. It should be noted (table 3) that the spreading of the tobacco mosaic virus protein on solutions of ammonium sulfate is a function of salt concentration only in the lower ranges, and that, in the cases of the more concentrated salt solutions, the area of the film obtained is essentially a constant; this behavior parallels that of egg albumin. This suggests a definite stable arrange-

ment of the tobacco mosaic virus protein molecules at the air-solution interface.

Electrophoretic behavior

The first electrophoretic studies on tobacco mosaic virus were made by Takahashi and Rawlins (47) in experiments on extracts of plants diseased with tobacco mosaic. They followed the migration of the virus by means of infectivity measurements and reported the material as being isoelectric below pH 4. Best (8), from precipitation studies, has reported the isoelectric point to be at pH 3.4. Electrophoretic measurements of chemically isolated tobacco mosaic virus protein dissolved in acetate buffers of ionic strength of about 0.02 were made by Eriksson-Quensel and Svedberg, using the Tiselius apparatus (14). The mobility-pH curve was found to have a slope, $d\mu/dpH$, of 1.23μ per second per volt per centimeter per pH unit, and the protein was found to be isoelectric at pH 3.49. In spite of the fact that the protein sample studied by these workers was not homogeneous with respect to sedimentation velocity, it was found to be electrophoretically homogeneous. Using the microcataphoresis apparatus of Northrop and Kunitz and adjusting the pH with hydrochloric acid and sodium hydroxide, Loring and Stanley (28) found that the isoelectric point of suspensions of crystals of the tobacco mosaic virus protein prepared by adding a small amount of ammonium sulfate to the solutions of the protein varied from pH 3.2 to 3.35. When acetate buffer was added the values were raised from 0.1 to 0.3 of a pH unit, depending upon the concentration of salts present.

Changes due to hydrogen-ion effects

In addition to the electrophoretic velocity which has already been discussed, there are other properties of the protein which vary with changes in the hydrogen-ion concentration. Sedimentation studies have been carried out by Eriksson-Quensel and Svedberg (14) and by Wyckoff (54) at various hydrogen-ion concentrations covering a wide range of the pH scale. It was found that, with the exception of the isoelectric zone, the sedimentation rate of the protein is essentially constant between pH 1.8 and 9.0, and that at hydrogen-ion concentrations greater or lower than those within this range the protein is rapidly broken down into components with lower sedimentation rates. It is interesting and important to note that this zone of approximately constant sedimentation rate coincides almost exactly with the pH zone of stability of the protein as measured by biological activity (figure 7).

Studies on the change of viscosity induced by altering the hydrogen-ion concentration have been made by Stanley (43) and by Lauffer (23). The

results of such a study are presented in figure 8. Double refraction of flow as measured by the method of Lauffer and Stanley (25) shows a behavior in general, though not rigorous, agreement with this viscosity behavior (figure 9) (23). These several behaviors suggest that, as one approaches the isoelectric point from either side, one encounters aggregation of the rod-shaped particles, at first end to end aggregation predominating and

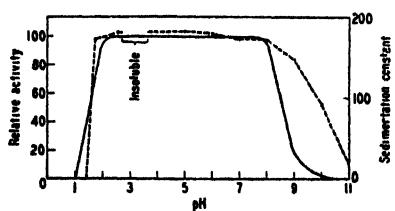


FIG. 7

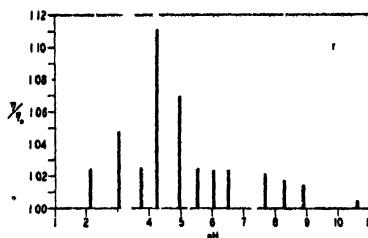


FIG. 8

Fig. 7. pH stability range of tobacco mosaic virus protein as measured by virus activity (solid line) and by sedimentation constant (dotted line). (Drawn from data of Best and Samuel, Stanley, Eriksson-Quensel and Svedberg, and Wyckoff, by Stanley (44)).

Fig. 8. The relative viscosity of solutions of tobacco mosaic virus protein (0.5 mg. per cubic centimeter) in buffers of ionic strength 0.02 at various pH values. (Drawn from data of Lauffer (23)).

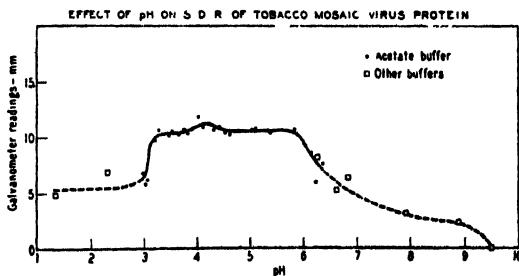


Fig. 9. Effect of pH changes on the stream double refraction of tobacco mosaic virus protein. (From Lauffer (23)).

later side to side aggregation predominating. This suggestion is in agreement with the observation that at the isoelectric point the protein exists in the form of needle-shaped crystals containing rod-shaped molecules packed side by side and end to end (7). Furthermore, the data of Mehl (31) show that the rotational diffusion constant evaluated from measurements of the orientation of tobacco mosaic virus particles in a mechanical field falls from 25 at pH 6.8 to 0.75 at pH 4.5. This is a direct indication

of a very considerable increase in length of the particle, substantiating the conclusion drawn from the viscosity and double refraction data.

REFERENCES

- (1) ALLARD: J. Agr. Research **6**, 649 (1916).
- (2) BAWDEN AND PIRIE: Proc. Roy. Soc. (London) **B123**, 274 (1937).
- (3) BAWDEN, PIRIE, BERNAL, AND FANKUCHEN: Nature **138**, 1051 (1936).
- (4) BECHHELD AND SCHLESINGER: Phytopath. Z. **6**, 627 (1933).
- (5) BEIJERINCK: Verhandel. Akad. Wetenschappen Amster. **6**, n.s., 1 (1898).
- (6) BERNAL: Proc. Roy. Soc. (London) **B125**, 299 (1938).
- (7) BERNAL AND FANKUCHEN: Nature **139**, 923 (1937).
- (8) BEST: Australian J. Exptl. Biol. Med. Sci. **14**, 1 (1936).
- (9) BOEDER: Z. Physik **75**, 258 (1932); J. Rheol. **3**, 494 (1932).
- (10) DUGGAR AND HOLLAENDER: J. Bact. **27**, 219, 241 (1934); Proc. Natl. Acad. Sci. U. S. **22**, 19 (1936).
- (11) DUGGAR AND KARRER [ARMSTRONG]: Ann. Missouri Botan. Garden **8**, 343 (1921); **10**, 191 (1923).
- (12) EISENSCHITZ: Z. physik. Chem. **A163**, 133 (1933).
- (13) ELFORD: Proc. Roy. Soc. (London) **B112**, 384 (1933).
- (14) ERIKSSON-QUENSEL AND SVEDBERG: J. Am. Chem. Soc. **58**, 1863 (1936).
- (15) FRAMPTON AND NEURATH: Science **87**, 468 (1938).
- (15a) FRAMPTON AND SAUM: Science **89**, 84 (1939).
- (16) GUTH: Kolloid-Z. **74**, 147 (1936).
- (17) GORTER: Proc. Acad. Sci. Amsterdam **37**, 20 (1934).
- (18) HERZOG, ILLIG, AND KUDAR: Z. physik. Chem. **A167**, 329 (1933).
- (18a) HILLS AND VINSON: Missouri Agr. Expt. Sta. Res. Bull. **286** (1938).
- (19) IWANOWSKI: Bull. Akad. Imp. Sci. St. Petersburg, n.s., III, **35**, 67 (1892).
- (20) KUHN: Kolloid-Z. **62**, 269 (1933).
- (20a) LANGMUIR AND SCHAEFER: J. Am. Chem. Soc. **59**, 1406 (1937).
- (21) LAUFFER: Science **87**, 469 (1938).
- (22) LAUFFER: J. Phys. Chem. **42**, 935 (1938).
- (23) LAUFFER: J. Biol. Chem. **126**, 443 (1938).
- (24) LAUFFER: Unpublished data.
- (25) LAUFFER AND STANLEY: J. Biol. Chem. **123**, 507 (1938).
- (26) LAVIN AND STANLEY: J. Biol. Chem. **118**, 269 (1937).
- (27) LORING: J. Biol. Chem. **128** (Sci. Proc. XXXII), lxxvi (1938).
- (28) LORING AND STANLEY: J. Biol. Chem. **117**, 733 (1937).
- (29) LORING, LAUFFER, AND STANLEY: Nature **142**, 841 (1938).
- (30) MACCLEMENT AND SMITH: Nature **130**, 129 (1932).
- (31) MEHL: Cold Spring Harbor Symposia Quant. Biol. **6**, 218 (1938).
- (32) NEURATH: Cold Spring Harbor Symposia Quant. Biol. **6**, 196 (1938).
- (33) NEURATH AND SAUM: J. Biol. Chem. **126**, 435 (1938).
- (34) PERRIN: J. phys. radium **7**, 1 (1936).
- (35) POLSON, A.: Nature **137**, 740 (1938).
- (36) ROSS AND STANLEY: Proc. Soc. Exptl. Biol. Med. **38**, 260 (1938); J. Gen. Physiol. **22**, 165 (1938).
- (37) SEASTONE: J. Gen. Physiol. **21**, 621 (1938).
- (38) SEASTONE: Private communication, July 8, 1938.
- (39) SMITH AND MACCLEMENT: Proc. Roy. Soc. (London) **B125**, 295 (1938).
- (40) STANLEY: Science **81**, 644 (1935); Phytopath. **26**, 305 (1936).

- (41) STANLEY: Am. J. Botany **24**, 59 (1937).
- (42) STANLEY: Ergeb. Physiol. biol. Chem. exptl. Pharmakol. **39**, 294 (1937).
- (43) STANLEY: J. Phys. Chem. **42**, 55 (1938).
- (44) STANLEY: Harvey Lectures 1937-1938, 170 (1938).
- (45) STANLEY AND LORING: Cold Spring Harbor Symposia Quant. Biol. **6**, 341 (1938).
- (46) SVEDBERG: Chem. Rev. **14**, 1 (1934); **20**, 81 (1937).
- (47) TAKAHASHI AND RAWLINS: Hilgardia **4**, 441 (1930).
- (48) TAKAHASHI AND RAWLINS: Proc. Soc. Exptl. Biol. Med. **30**, 155 (1932); Science **77**, 26, 284 (1933).
- (49) TAKAHASHI AND RAWLINS: Science **85**, 103 (1937).
- (50) THORNBERRY: Phytopath. **25**, 36, 601, 938 (1935).
- (51) WAUGH AND VINSON: Phytopath. **22**, 29 (1932).
- (52) WIENER: Abhandl. math.-phys. Klasse sachs. Akad. Wiss. **32**, 507 (1912).
- (53) WYCKOFF: J. Biol. Chem. **121**, 219 (1937).
- (54) WYCKOFF: J. Biol. Chem. **122**, 239 (1937).
- (55) WYCKOFF AND COREY: J. Biol. Chem. **118**, 51 (1936); Science **84**, 513 (1936).
- (56) WYCKOFF, BISCOE, AND STANLEY: J. Biol. Chem. **117**, 57 (1937).
- (57) ZOCHER AND JACOBSSON: Kolloid-Z. **41**, 220 (1927).

CHEMICAL ASPECTS OF THE PRECIPITIN AND AGGLUTININ REACTIONS¹

MICHAEL HEIDELBERGER

*Department of Medicine, College of Physicians and Surgeons, Columbia University
and the Presbyterian Hospital, New York, New York*

Received January 30, 1939

Ten years ago the inclusion of studies on precipitin and agglutinin reactions in a symposium on the physical chemistry of the proteins would have occasioned raised eyebrows. To-day, there may still be some tendency to produce this effect, but I hope soon to have demonstrated that the subject properly belongs in such a symposium. In doing so it will be necessary to link some of the terms of immunology with those of chemistry, and you will therefore, I trust, pardon the introduction of a few definitions.

When a foreign protein, called an *antigen*, is injected into an animal, there usually follows, some time later, the appearance of new substances in the blood serum of the animal. These substances, called *antibodies*, are characterized by their property of reacting with the antigen injected. When this interaction results in the formation of a precipitate it is called a "precipitin reaction." Antigen-antibody interaction is *specific*, that is, antibodies arising from the stimulus of a given antigen, such as crystalline egg albumin, do not precipitate solutions of another albumin, such as crystalline horse serum albumin. More closely related proteins may, however, cause animals to produce antibodies which overlap in part, and reactions between such antigens and antibodies evoked by closely related antigens are called "cross reactions."

An invading microorganism may, for present purposes, be considered as a collection of antigens. If these antigens stimulate the production of antibodies, these may interact with the invader, causing its direct or indirect destruction, and the animal is said to be *immune*. Antibodies to a microorganism may react not only with antigens extracted from the invading cell to give precipitin reactions, but may react with such components on the effective bacterial surface. The bacteria then clump together, and

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

this special case of the precipitin reaction at the reactive microbial surfaces is called an "agglutinin reaction."

In the large group of encapsulated microorganisms to which the pneumococcus, or germ of pneumonia, belongs, the dominant cellular antigen occurs in the capsular layer. The complete antigen has not yet been isolated, but each serological type of pneumococcus (and there are some forty) appears to be characterized by a chemically distinct polysaccharide to which these type specific serological reactions are due.² When isolated, the carbohydrates are found to be capable of entering into precipitin reactions with antisera from animals injected with the same pneumococcus type, and these antisera also agglutinate pneumococci of the homologous

TABLE I
Properties of immunologically specific polysaccharides of pneumococcus

| TYPE | NITROGEN per cent | $[\alpha]_D$ | NEUTRAL EQUIVALENT | ACETYL per cent | URONIC ANHYDRIDE per cent | REDUCING SUGARS AFTER HYDROLYSIS per cent | η_{relative} 0.2 PER CENT IN H_2O | η_{relative} 0.1 PER CENT IN 0.0 PER CENT NaCl |
|-------|----------------------|--------------|-----------------------|--------------------|---------------------------------|-------------------------------------------------------|----------------------------------------------------------|------------------------------------------------------------------------|
| I | 4.6* | +280° | 650 | 7.1† | 56 | 30 | 9 | 1.7 |
| II | 0.2 | +55° | 950 | 1 | 18 | 95 | 11 | 1.6 |
| III | 0.1 | -35° | 340 | 0.5 | 50 | 85 | 32 | 3.0 |
| IV | 5.5 | +30° | 1500 | 5.8‡ | — | 71 | — | — |
| VIII§ | 0.2 | +123° | 720 | 0.5 | 27 | 87 | 17 | 2.5 |

* About one-half reacts as amino nitrogen.

† Reacts as *O*-acetyl (Avery, O. T., and Goebel, W. F.: J. Exptl. Med. **58**, 731 (1933); Pappenheimer, A. M., Jr., and Enders, J. F.: Proc. Soc. Exptl. Biol. Med. **31**, 37 (1933); Enders, J. F., and Wu, C. J.: J. Exptl. Med. **60**, 137 (1934)).

‡ Reacts as *N*-acetyl (Heidelberger, M., and Kendall, F. E.: J. Exptl. Med. **53**, 625 (1931)).

§ Goebel, W. F.: J. Biol. Chem. **110**, 391 (1935); Heidelberger, M., Kabat, E. A., and Shrivastava, D. L.: J. Exptl. Med. **65**, 487 (1937).

type, owing to interaction with specific polysaccharide on the bacterial surfaces. Since these specific polysaccharides have functioned as "antigen" in chemical studies on the precipitin reaction, the properties of the most thoroughly studied members of the series are given in table 1.

Since our discussion will deal only in part with protein antigens, let us review briefly the present status of our knowledge of the chemical nature of antibodies. It has long been known that antibodies occur in one or more of the globulin fractions of serum. Felton (4) has shown that, when antipneumococcus horse sera are poured into slightly acidified water, as

* For reviews see references 9 and 1.

much as 90 per cent of the serum proteins may remain in solution, but 60 to 80 per cent of the antibodies may precipitate. After removal of inactive material with acid and treatment of the residual antibody protein with zinc or aluminum salts, Felton (5, 6) isolated metal-antibody complexes which were completely precipitable by the pneumococcus polysaccharide of homologous type. Unfortunately, the globulins remaining after removal of the zinc or aluminum ions were only 80 per cent specifically precipitable. As a consequence, however, of the studies on the mechanism of the precipitin reaction about to be described, it has been possible to isolate from antipneumococcus sera antibody globulin solutions in which up to 100 per cent of the protein present was specifically reactive. Analytically pure antibody has been prepared in this way in two laboratories. It has the properties of serum globulin, as will appear later. If the protein in such solutions were adsorbed on a small amount of true antibody of unknown chemical nature, it would be a qualitatively distinct protein fraction known to be absent in a normal serum which does not precipitate antigen, and since this indirectly defines the term "antibody," antibody is protein. Additional evidence for this conclusion will be discussed below.

The early theory of Ehrlich that immune reactions such as the precipitin reaction involved the chemical combination of antigen and antibody in stoichiometrical proportions was opposed by that of Bordet, which held them to be merely adsorption phenomena. It was soon necessary to modify the latter view in order to account for the specificity of immune reactions, and an initial chemical reaction was gradually assumed, followed, in the precipitin and agglutinin reactions, by flocculation due to the presence of electrolytes. Arrhenius and Madsen stressed the reversibility of some antigen-antibody reactions, applied the law of mass action, and found analogies with the reaction of weak acids and weak bases.³ It must be remembered that at the time these theories were enunciated, little was known of the chemical nature of either antigens or antibodies, and the only analytical methods available were either of biological nature or the essentially qualitative and entirely relative serological dilution methods with their large capacities for error.

Before undertaking studies on the actual mechanism of the precipitin reaction, Dr. Kendall and I therefore found it necessary to abandon such techniques and to devise absolute micromethods for the estimation of antigen and antibody, conforming to the criteria of analytical chemistry and yielding accurate data regardless of the presence of non-specific protein.⁴ Precipitates formed from accurately measured quantities of anti-serum and antigen dilutions were washed under definite conditions with

³ For a more detailed discussion see reference 44.

⁴ Reviewed by Heidelberger (10); see also reference 19.

0.9 per cent sodium chloride solution and analyzed for nitrogen. In the proper regions of the precipitin reaction range this procedure permits the estimation of antigen or antibody in milligrams per milliliter.

The first instance of the precipitin reaction studied in this way by Kendall and myself was that between the nitrogen-free specific polysaccharide of Type III pneumococcus^a and the partially purified antibody in a Felton solution prepared from Type III antipneumococcus horse serum. This choice of "antigen" greatly simplified the analytical problem of the composition of the precipitates, for the polysaccharide made no contribution to the nitrogen precipitated. Once it had been shown that non-specific nitrogen was not involved, the nitrogen estimations became a direct measure of the amount of antibody precipitated.

If a very small amount of S III is added to a relatively large amount of antibody, A, it is found that more than 40 mg. of antibody nitrogen may be precipitated for each milligram of S III. If increasing amounts of S III are added in separate portions to the same quantity of A, this ratio in the precipitate decreases steadily, with no evidence of discontinuity. In this region of the reaction range no S III can be found in the supernatants by the delicate serological test sensitive to a dilution of 1:10,000,000, so that it is assumed that all of the S III added is in the precipitate. In this region antibody is still in excess, as may be shown by addition of a little S III to a portion of the supernatant. With still larger amounts of S III a region of the reaction range follows in which neither S III nor A is demonstrable in the supernatant from the precipitate, and this we have termed the "equivalence zone." With still larger amounts of S III the polysaccharide finally appears in the supernatant, and in this region precipitation of antibody is at a maximum. When still greater quantities of S III are added, antibody nitrogen precipitated remains constant while more of the S III added enters into combination. Finally, in some antibody solutions, at least, the precipitate attains constant composition, and with larger amounts of S III, less and less precipitate is formed, until finally precipitation is entirely inhibited. The region of diminishing precipitation is therefore termed the "inhibition zone." The course of the reaction is illustrated by the curves in figure 1, except for the inhibition zone, and it was found at an early stage of the work that such curves, up to the equivalence zone, could be expressed by the empirical equation

$$\text{Milligrams of antibody nitrogen precipitated} = aS - bS^2 \quad (\text{I})$$

In a sense these reactions are reversible, for the precipitate formed in the region of excess antibody takes up S III when shaken with a solution of the

^a Subsequently referred to as S III. These studies were carried out with preparations now known to have been degraded by heat (24). While this would affect the numerical values of the results, it in no way modifies the principles involved.

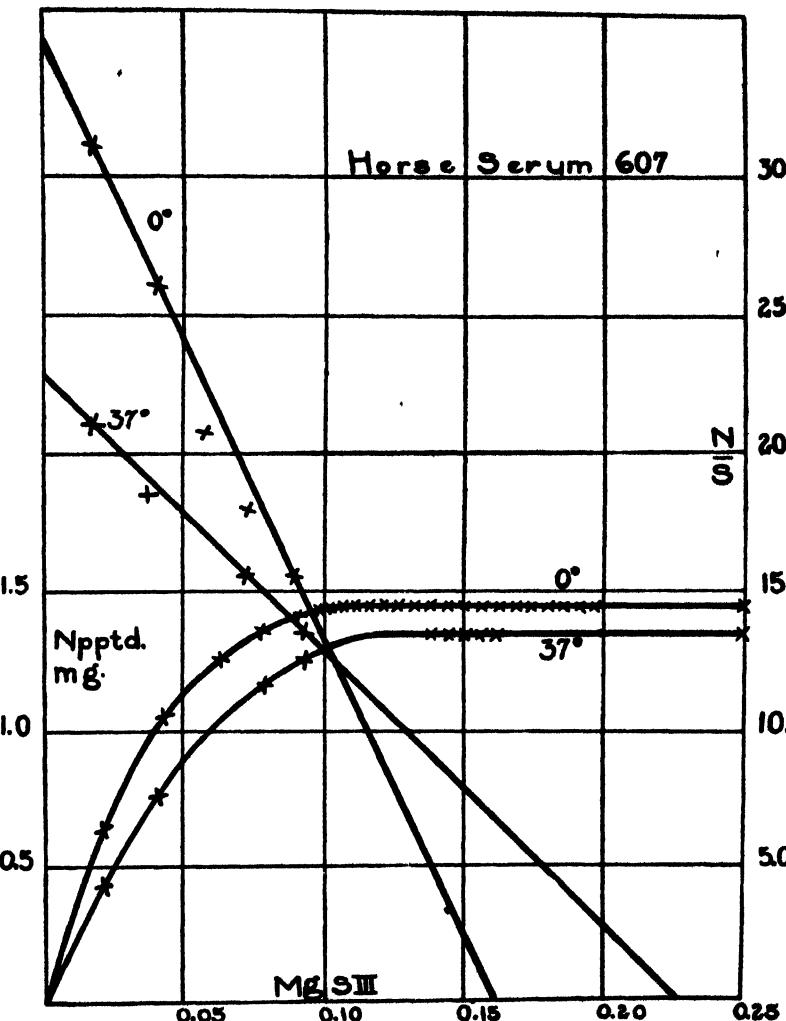
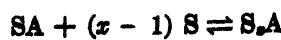


FIG. 1

polysaccharide and even dissolves in relatively concentrated S III solutions. The three limiting regions of the reaction range may therefore be represented by



in the first two of which equilibrium must be far to the right, since measurable dissociation cannot be detected. In these equations SA_4 represents in arbitrary units, not molecules, the limiting compound formed in the region of excess antibody, SA the composition of the precipitate at the midpoint of the equivalence zone, and S_xA the composition of the soluble compound or compounds in the inhibition zone. It could be shown that S_xA contained one more molecule of S III than the precipitate in the inhibition zone in which it was in apparent equilibrium (16), in confirmation of the belief of Arrhenius and others in the presence of a soluble antigen-antibody compound in this zone.

Since antibody probably exists in solution in physiological media as an ionized globulin-sodium chloride complex, and S III is the highly ionized salt of a polyvalent aldobionic acid (12, 17), the reactions are probably

TABLE 2*
Effect of volume, final concentration of antibody, and time of standing

| VOLUME cc. | ANTIBODY B 62 AT 0°C. | | | | ANTIBODY B 61 AT 37°C. | | | |
|---------------|--------------------------------------------------------------------|--------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------|-------------------------------------------------------|--------------------------------------------------------------|-------------------------------------------------------|--------------------------------------------------------------|
| | Antibody nitrogen pptd. by 0.08 mg. S III in 24 hr. | Concen- tration of antibody nitrogen mg. per cc. | Antibody nitrogen pptd. by 0.08 mg. S III in 48 hr. | Concen- tration of antibody nitrogen mg. per cc. | Antibody nitrogen pptd. by 0.08 mg. S III | Concen- tration of antibody nitrogen mg. per cc. | Antibody nitrogen pptd. by 0.10 mg. S III | Concen- tration of antibody nitrogen mg. per cc. |
| 2 | 0.88 | 0.21 | 0.87 | 0.21 | | | | |
| 4 | 0.84 | 0.11 | 0.91 | 0.10 | 0.87 | 0.25 | 1.15 | 0.18 |
| 6 | 0.83 | 0.08 | 0.87 | 0.07 | | | | |
| 8 | 0.88 | 0.05 | 0.84 | 0.06 | 0.87 | 0.12 | 1.16 | 0.09 |
| 10 | 0.84 | 0.05 | 0.84 | 0.05 | | | | |
| 12 | 0.82 | 0.04 | 0.87 | 0.04 | 0.85 | 0.08 | 1.16 | 0.06 |

* Reprinted from reference 20.

ionic, and the application of the mass law in some form would seem justified. The precipitin reaction between S III and homologous antibodies would then be merely a more complex instance of a specific precipitation, such as that between barium and sulfate ions. Even the inhibition zone may have at least a partial analogy in the well-known solubility of silver cyanide in excess cyanide solution.

There are, however, difficulties in the way of a simple quantitative formulation of the reaction in terms of the law of mass action. In the first place, there is no evidence of the existence of definite stages between the limit of SA_4 in the region of large antibody excess and SA in the equivalence zone, or between SA and $S_{x-1}A$. This might be overcome by assuming a continuous series of solid solutions, or that the mutual multivalence of S and A is so great that a continuous series of compounds could result.

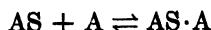
While there is both structural (12, 17) and other (18) evidence that S III contains a number of immunologically reactive groupings, or valences, and while as much as is known of the structure of proteins permits the assumption of recurrent groups of amino acids which might be the centers of specific combination, there are valid objections to these views. Foremost is the finding that the composition of the precipitate depends not upon the antibody concentration at equilibrium (20) (table 2) but on the proportions in which the components are mixed. This curious state of affairs not only prevents a simple treatment of the precipitin reaction according to the law of mass action, but also blasts the hopes of those colloid enthusiasts who have endeavored to characterize this and other immune reactions by adsorption isotherms, for adsorption isotherms, too, contain a concentration term.

With the aid of several assumptions, a relation may still be found which accounts quantitatively for this and many other instances of the precipitin reaction, and may, moreover, be derived from the law of mass action. This is possible if the reaction be considered as a series of successive bimolecular reactions which take place before precipitation occurs (20). The first step in the reaction between S III and A would then be



Since both S III and antibody may be considered multivalent⁶ with respect to each other, as explained above, the SA formed in this reaction could react with other molecules of the same compound, or with S III, or with A, whichever is present in excess.

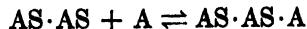
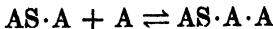
Let us consider the subsequent course of the reaction in the region of excess antibody. The second step would consist of the two competing bimolecular reactions:



and



There would follow a third step, in which the competing bimolecular reactions involved would be



and



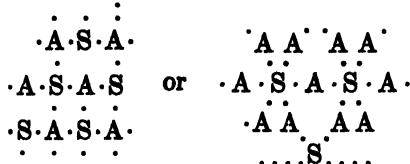
⁶ That is, with a valence of 2 or more.

in which the first two reactions would occur only in the presence of enough A to carry the composition of the reaction product beyond the A_2S stage. Similarly, each compound formed in the third step would react with each other compound, or with more A, if present, to form still more complex substances, and the reaction would continue until particles would be formed large enough to settle from the solution, and precipitation would take place. The mutual discharge, with loss of affinity for water, of ionized groupings brought together by the series of chemical reactions would doubtless facilitate precipitation (32).

If A and S III are mixed in equivalent proportions, the AS formed in reaction 1 would merely polymerize in steps 2, 3, etc., and the equivalence point precipitate would be $(AS)_n$.

In the region of excess S III a similar series of expressions would apply, in which S and A would be interchanged in steps 2, 3, etc. In the presence of a large excess of S, in the inhibition zone, there would also be present in solution a soluble compound, AS_n , containing one more molecule of S in combination than the last insoluble compound (16). Since this is formed only with a very large excess of S, all of the specific groupings of A would tend to react with S rather than with AS complexes and there would be no large, insoluble, intermolecular aggregates formed.

The final precipitate, then, would in each case consist of antibody molecules held together by varying proportions of S III molecules, thus:



This view is similar to that presented recently by Marrack (32) but, it is believed, is more definite and more easily treated quantitatively. The process of aggregation, as well as the initial hapten-antibody combination, is considered to be a chemical reaction between definite molecular groupings.

In the mathematical treatment it is assumed, first, that the antibody, which is known to be a mixture of antibodies of varying reactivities, may be treated as a single substance; second, that in the initial stage of the reaction A reacts with S to give only AS; third, that in the second step of the reaction the products are $AS \cdot A$ and $AS \cdot AS$; fourth, that the mass law applies, so that the rates of formation of $AS \cdot A$ and $AS \cdot AS$ are proportional to the concentrations of the reacting substances; and fifth, that the dissociation of $AS \cdot A$ and $AS \cdot AS$ is negligible. Although there is no

reason to assume discontinuities in the building up of the final aggregates, the reactions are arbitrarily treated as successive stages in order to simplify the mathematics involved.

At the beginning of the second stage of the reaction, then, in the presence of excess antibody, let

A = total units of antibody in the reacting system,

B = units of AS formed in the first step = units of S added,

$A - B$ = units of free antibody at end of first step,

x = units of AS·A formed at time t ,

y = units of AS·AS formed at time t , and

V = volume.

Then

$$\frac{A - B - x}{V} \quad \text{concentration of free antibody at time } t$$

and

$$\frac{B - x - 2y}{V} \quad \text{concentration of AS at time } t$$

$$\text{Rate of formation of AS·A} = \frac{dx}{dt} = K \left(\frac{A - B - x}{V} \right) \left(\frac{B - x - 2y}{V} \right) \quad (4)$$

$$\text{Rate of formation of AS·AS} = \frac{dy}{dt} = K' \left(\frac{B - x - 2y}{V} \right)^2 \quad (5)$$

Dividing equation 4 by equation 5,

$$\frac{dx}{dy} = \frac{A - B - x}{B - x - 2y}$$

if $K = K'$. Integrating,

$$\frac{A - 2x}{2(A - B - x)^2} = \frac{y}{(A - B - x)^2} + C$$

To evaluate C : at start of reaction when $t = 0$, $x = 0$, $y = 0$,

$$C = \frac{A}{2(A - B)^2}$$

At the end of the reaction $x + 2y = B$. Therefore,

$$x = \frac{AB - B^2}{A} \quad \text{and} \quad y = \frac{B^2}{2A}$$

Since each unit of x and y contains 2 units of antibody, the number of units of antibody precipitated is given by $2(x + y)$ which equals

$$2B - \frac{B^2}{A} \quad (6)$$

It will be noted that the volume factors cancel, so that the amount of antibody precipitated depends only on the relative amounts of antibody and S III present and not on their concentrations.

This treatment of the problem involves only the formation of compounds having ratios between R and $2R$, where R is the ratio of antibody to S III in the equivalence point compound. The experimental data show that compounds having ratios greater than $2R$ may be formed, for at 0°C. in the presence of a large excess of antibody ratios greater than $4R$ are encountered. By extending the process used for the calculation of the second step to stage 3 and beyond, it is possible to calculate the amount of antibody precipitated by a given amount of S III when the ratio varies between R and $3R$ and also between R and $4R$.

Similar formulas apply in the region of excess S, and in their derivation S and A are interchanged. In applying the equations to experimental data it is necessary to convert units into milligrams. This may be done by defining 1 unit of antibody nitrogen as 1 mg. A may then be put equal to the number of milligrams of antibody nitrogen precipitated at the equivalence point and R equal to the ratio of A to milligrams of S III at this point. Equation 6 then becomes

$$\text{Milligrams of antibody nitrogen precipitated} = 2RS - \frac{R^2S^2}{A} \quad (7)$$

If the validity of these considerations be admitted, the theoretical significance of a and b in the empirical equation I is now clear, for $a = 2R$ and $b = R^2/A$. Both of these constants have the immunological and chemical significance stated above. If, instead of taking the reference point as the difficultly determinable "equivalence point," it be taken as either end of the equivalence zone, depending on the precipitin system studied, it will usually be found possible to avoid the more complicated R to $3R$ equations given in reference 20 and use only equation 7. The application of this to a number of antibody solutions and sera is shown in table 3. The agreement between found and calculated values is very close.

If both sides of equation 7 be divided by S, the resulting equation,

$$\frac{N}{S} = 2R - \frac{R^2}{A} S$$

is that of a straight line. This linear relationship makes it possible to characterize an unknown Type III antipneumococcus serum or antibody solution in the region of excess antibody by two analyses, in duplicate. If the ratio of antibody nitrogen (N) to S III precipitated be determined for two different amounts of S III in the region of excess antibody and a

straight line be drawn through the two points so obtained, the intercept on the y -axis = $2R$ and the slope = $-R^2/A$. With the R and A values at the beginning of the equivalence zone calculated in this way, the amount of antibody nitrogen precipitated by any quantity of S III less than A/R may be calculated with a fair degree of accuracy.

TABLE 3*

Comparison of experimental data with values calculated according to

$$\text{nitrogen precipitated} \quad 2RS - \frac{R^2S^2}{A}$$

| ANTIBODY NO..... | BVA | | B 36 | | B 62 | | B 61 | | B 62 | | Serum 607 | | Serum 607 | |
|---------------------|-------------------|--------|-------------------|--------|-------------------|--------|-------------------|--------|-------------------|--------|-------------------|--------|-------------------|--------|
| TEMPERATURE, °C.... | 37 | 0 | 37 | 0 | 37 | 37 | 0 | 0 | 37 | 37 | 0 | 0 | 37 | 37 |
| R..... | 18.6 | | 12.4 | | 11.4 | | (17)† | | 12 | | (15) | | (11) | |
| A..... | 4.08 | | 1.86 | | 1.71 | | (1.23)† | | 1.20 | | (1.42) | | (1.81) | |
| S III used | Nitrogen pptd. | |
| | Found | Calcd. |
| mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. | mg. |
| 0.01 | | | | | | | 0.36 | 0.32 | | | | | | |
| 0.02 | | | 0.50 | 0.46 | 0.45 | 0.43 | 0.57 | 0.59 | 0.44 | 0.43 | 0.62 | 0.54 | 0.42 | 0.40 |
| 0.03 | | | | | | | 0.78 | 0.81 | | | | | | |
| 0.04 | | | | | 0.79 | 0.79 | | | | | 1.03 | 0.95 | 0.74 | 0.73 |
| 0.05 | 1.22 | 1.25 | 1.03 | 1.03 | 0.97 | 0.95 | 1.07 | 1.11 | | | | | | |
| 0.06 | | | | | 1.08 | 1.09 | | | 0.96 | 1.01 | 1.25 | 1.23 | | |
| 0.075 | | | 1.41 | 1.40 | | | | | | | 1.35 | 1.36 | 1.16 | 1.13 |
| 0.08 | | | | | 1.29 | 1.34 | | | | | | | | |
| 0.09 | | | | | | | | | | | 1.40 | 1.42 | 1.23 | 1.23 |
| 0.10 | 2.24 | 2.27 | 1.66 | 1.65 | 1.54 | 1.52 | | | | | | | | |
| 0.12 | | | | | 1.68 | 1.64 | | | | | | | | |
| 0.20 | 3.62 | 3.62 | | | | | | | | | | | | |
| 0.25 | 3.87 | 3.96 | | | | | | | | | | | | |

* Reprinted from reference 20.

† R and A values in parentheses deduced from nearest actual determination.

In employing equation 7 in the region of excess antibody, it is assumed that all of the antibody present is precipitated at the beginning of the equivalence zone. This is, however, not actually the case, for the amount of antibody nitrogen precipitated does not reach a maximum until S III is present in appreciable excess. For the complete description of the behavior of a serum in the precipitin reaction a separate determination of the maximum amount of specifically precipitable nitrogen is necessary.

In the region of excess S III the behavior of a serum as far as the begin-

ning of the inhibition zone may be characterized by the determination of the A and S III precipitated at two points, since in this region the relation, linear with respect to 1/Total S,

$$\frac{S_{\text{pptd.}}}{A} = 2R' - \frac{(R')^2 A}{\text{Total } S}$$

applies if R' be taken as the S/A ratio at the end of the equivalence zone at which S III appears in excess and A be taken as the amount of antibody precipitated.

In the inhibition zone, in which large amounts of S III are present and the amount of precipitate has begun to diminish, this equation is no longer applicable, and it is necessary to determine the apparent dissociation constant of the soluble compound S_sA (cf. also 21).

In spite of the wide variation in the properties of individual sera, these expressions permit the complete description of the behavior of an unknown antiserum with S III without an unduly burdensome number of microanalyses or the sacrifice of a large amount of material.

Application of these principles to precipitin systems in which the antigen is a protein is complicated by the necessity of distinguishing between antigen and antibody nitrogen if the composition of the precipitate is to be determined. This was accomplished with the aid of a red protein dye, R-salt-azobiphenylazo-crystalline egg albumin, which was freed from fractions reactive in anti-egg albumin sera and then injected into rabbits. In this antigen-antibody system antigen could be estimated colorimetrically in an alkaline solution of the washed specific precipitate, while antibody was determined by difference after a total nitrogen estimation. Equation 7 was found to be applicable in this system as well (21). With the aid of the information gained, a study was made of a colorless, individual protein, crystalline egg albumin, and its homologous antibodies, and this instance of the precipitin reaction was also found to be quantitatively described over a large part of the reaction range by the theory (22). The precipitin systems with crystalline horse serum albumin (30) and various thyroglobulins (38) and their homologous antibodies were also found to conform. The latter work afforded an opportunity of testing the assumption, made with S III and with egg albumin, that in the absence of a positive test in the supernatant all of the presumably pure antigen added in the region of excess antibody would be in the precipitate, and, in the latter instance, could be deducted in order to calculate the amount of antibody precipitated. This assumption has been criticized by Taylor, Adair, and Adair (39), but was found justified in the one instance in which a direct test was possible. From 96 to 101 per cent of the thyroglobulin iodine added was precipitated in this region. Very recently, Pennell and

Huddleson (36) have found the above quantitative precipitin theory applicable to systems involving bacterial antigens of the *Brucella* group.

It is also possible, as Kabat and I have shown, to adapt the microanalytical methods used for studying the precipitin reaction to the estimation of pneumococcus agglutinins (13) and with these methods to show that Type I pneumococcus and its homologous antibody undergo reaction according to the same type of equation (14). The reaction is, however, simpler, since the exigencies imposed by the reactive bacterial surfaces limit the range of combining proportions of bacterial polysaccharide, and antibody.

However, the mere agreement of even a large body of data with a theory does not demonstrate its validity. The microanalytical methods upon which the theory is based have met with general acceptance (cf., for example, 3, 28, 33, 36, 39) and have been extended to the toxin-antitoxin reaction (35), but the quantitative theory has fared less well (cf., for example, 29, 32). These objections will be discussed in detail on another occasion. It may be said at this time, however, that even though certain faults of the theory have been proclaimed repeatedly by Kendall and myself, and no more has been claimed for it than a first attempt at a general quantitative theory of two important immune reactions, it can explain much that is not accounted for by the older, essentially qualitative theories, and it has permitted several rather far-reaching predictions which could not otherwise have been foreseen.

For example, a study of the effect of strong salt solutions on the reaction between pneumococcus polysaccharide and homologous antibodies (25) showed that the diminished precipitation and the decrease in the values of both constants in equation 7 (table 4) was not due to increased solubility of the precipitate. On the basis of the quantitative theory a reversible shift in the equilibrium between polysaccharide and antibody was indicated, and the prediction was made that use could be made of this shift for the isolation of pure antibody. If, for example, 0.1 mg. of S III precipitated 1.24 mg. of antibody nitrogen from 1 ml. of a given antiserum in 0.15 M sodium chloride and only 1.01 mg. of nitrogen in 1.75 M sodium chloride, it should be possible, if the equilibria involved were reversible, to add 1.75 M sodium chloride to the washed precipitate formed in 0.15 M sodium chloride and dissociate 0.23 mg. of nitrogen, representing analytically pure antibody. Practical difficulties in washing and handling larger amounts of material prevented the immediate realization of this ideal, but solutions of 90 to 98 per cent purity were readily obtainable in a single step from many antipneumococcus horse and rabbit sera of various types (23). With refinements in technique, analytically pure antibody

globulin was isolated (15), after this had already been accomplished by other slight modifications of our method by Goodner and Horsfall (7).

Some of the highly purified antibody globulin solutions were studied in Upsala, Sweden, in the Svedberg ultracentrifuge. Pedersen and I found (26, 27) that the Type III pneumococcus antibody produced in the rabbit showed only a single component with a sedimentation constant of

TABLE 4*

Effect of varying sodium chloride concentrations on reaction between S III and antiserum 607 at 0°C. and at 37°C.

| S III USED | ANTIBODY NITROGEN PRECIPITATED AT 0°C. | | ANTIBODY NITROGEN PRECIPITATED AT 37°C. | |
|--------------------------------------------|----------------------------------------|--------------------------|-----------------------------------------|--------------------------|
| | 0.15 M NaCl | 1.75 M NaCl | 0.15 M NaCl | 1.75 M NaCl |
| mg. | mg. | mg. | mg. | mg. |
| 0.03 | | 0.48 | | 0.42 |
| 0.04 | 0.74 | | 0.64 | |
| 0.07 | | 0.83† | | 0.78 |
| 0.075 | 1.12 | 0.91 | 0.97 | |
| 0.10 | 1.24 | 1.01 | 1.09 | 0.87 |
| 0.15‡ | | 1.07 | | 0.91 |
| 0.30‡ | | 1.10 | | 0.94§ |
| 1.00‡ | | 1.08 | | 0.85 |
| 2.34‡ | | 0.87 | | 0.54 |
| EQUATIONS: | | | | |
| mg. of antibody nitrogen precipitated..... | 22.5 S-101 S ^a | 18.4 S-83 S ^a | 19.3 S-85 S ^a | 16.2 S-75 S ^a |
| A ¶..... | 1.25 | 1.02 | 1.10 | 0.87 |

* Reprinted from reference 25.

† The supernatant remained clear when dialyzed at 0°C. against 0.9 per cent saline. The nitrogen determinations were run according to Teorell (Acta Med. Scand. **68**, 305 (1928)).

‡ Excess SIII.

§ One determination.

¶ Calculated milligrams of antibody nitrogen precipitated at antibody excess end of equivalence zone.

7×10^{-13} , the same as that of the principal component of normal globulin (figure 2). A similar preparation from horse serum was also monodisperse, but showed a sedimentation constant of 18×10^{-13} , close to that of a minor globulin fraction of relatively high molecular weight present in all mammalian sera (figure 3). Removal of the cell contents at different levels after centrifugation showed that the antibody had actually sedimented with the heavy portion. This pointed to an entirely different mechanism

for the production of pneumococcus anticarbohydrate in the horse and the rabbit. Goodner, Horsfall, and Bauer, at the same time, concluded from ultrafiltration experiments that much of the anticarbohydrate in rabbit sera was smaller than that in horse sera (8).

The same antibodies produced in other species of animals were purified in the laboratory of the Presbyterian Hospital and were shown by Kabat

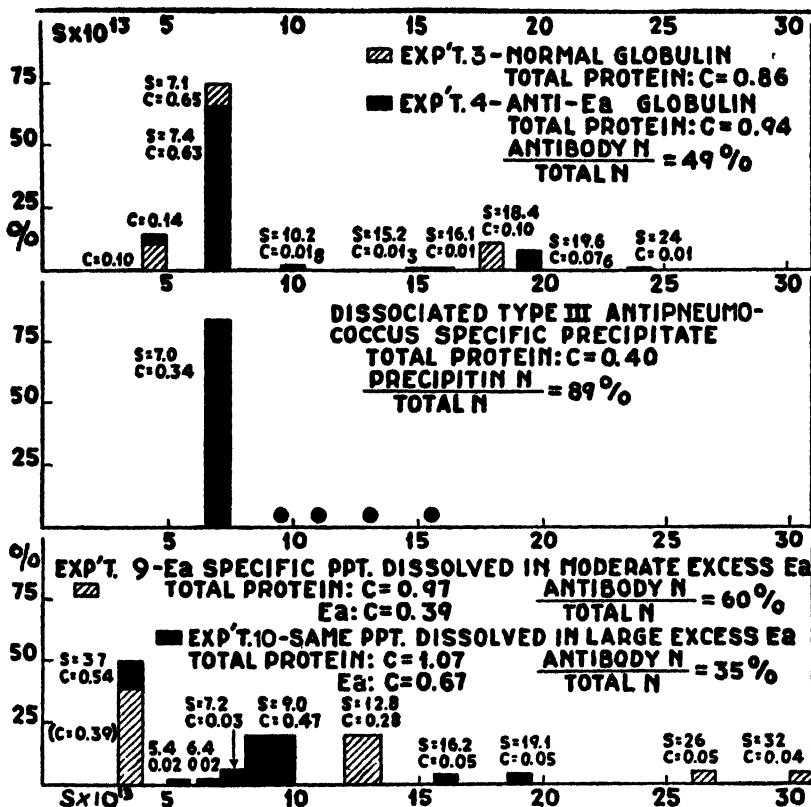


FIG. 2

and Pedersen (31) to exhibit similar differences. Table 5 shows the molecular weights established for the antibodies of the different species by these workers.

From the above studies it is clear that pneumococcus anticarbohydrate and anti-egg albumin engendered in the rabbit have the same molecular size as that of the principal component of normal serum globulin, while the former antibody, as produced by the horse, cow, and pig, has a molec-

ular weight six times as great, corresponding to a minor component also present in all normal sera. If there could any longer be doubt as to the actual protein nature of antibodies, it would be dispelled by the typical protein electrophoresis curves found for several of the above antibody solutions by Tiselius (40) and by the study of the electrophoretic behavior of antibodies in whole antisera made by Tiselius and Kabat (41).

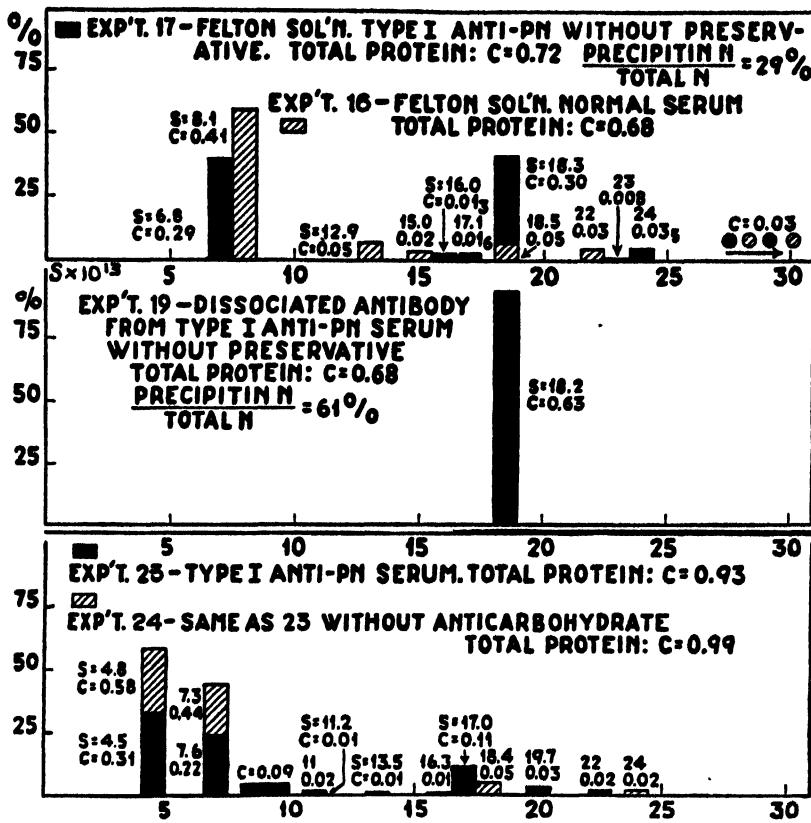


FIG. 3

The new information gained from this work has had its practical repercussions, as well, for it has furnished much of the theoretical background and practical methods of control for the use of antipneumococcus rabbit sera in the treatment of pneumonia (43).

One other feature of the sedimentation diagrams shown in figure 2 is of interest. The last two diagrams illustrate the behavior of egg albumin-anti-egg albumin precipitates dissolved in an excess of egg albumin solution.

Evidence has already been given for the existence of actual chemical compounds in the inhibition zone of the S III-antibody system, but the visualization of such compounds in the ultracentrifuge would seem to place their existence beyond question. It should be remembered that the specific precipitate was in this instance composed of aggregates of egg albumin, with a sedimentation constant of about 3.6×10^{-13} , and antibody with $s = 7 \times 10^{-13}$. The presence in the solution of a small number of components of higher sedimentation constant than either would seem to indicate the formation of definite compounds, for if the precipitate were merely "peptized" by the excess of antigen, as the colloidal explana-

TABLE 5*

Molecular weights of pneumococcus anticarbohydrate formed in animals of various species

| SPECIES | $S_20 \times 10^{13}$ | $D_20 \times 10^7$ | M | f/f_0 |
|-------------|-----------------------|--------------------|---------|---------|
| Pig..... | 18.0 | 1.64 | 980,000 | 2.0 |
| Cow..... | 18.1 | 1.69 | 910,000 | 2.0 |
| Horse..... | 17.9 | 1.63 | 980,000 | 2.0 |
| Rabbit..... | 7.0 | 4.23 | 157,000 | 1.4 |
| Monkey..... | 6.7 | 4.06 | 157,000 | 1.5 |

* Reprinted from reference 31.

S = sedimentation constant; D = diffusion constant; M = molecular weight; f/f_0 = dissymmetry constant.

Fig. 1. Precipitation of Type III antipneumococcus horse serum by Type III pneumococcus specific polysaccharide.

Fig. 2. Molecular species found in ultracentrifuge runs on rabbit sera and antibody solutions. (Reprinted from reference 26.)

Fig. 3. Molecular species found in ultracentrifuge runs on horse sera and antibody solutions. (Reprinted from reference 26.)

tion goes, there is no *a priori* reason why this should stop short of the actual reaction components themselves.

An additional result of the quantitative precipitin and ultracentrifugal studies is the possibility, now that the molecular weights of the reactants are known, of writing empirical formulas for the composition of specific precipitates at limiting points in the reaction range (11). The data are given in table 6 and need little comment, except that the formulas are to be considered as average compositions, and not those of single molecular species. However, if the conception of precipitate formation by the union of multivalent antigen with multivalent antibody is correct, each particle of precipitate is a molecule of vast size, of approximately the empirical composition indicated for a given point in the reaction range. Nor are

the compositions given of such extreme range as to render preposterous a treatment of the reaction according to the laws of classical chemistry.

Another prediction, made from the quantitative agglutinin theory, is of chemical interest, since its verification explains the rôle of salts in these immune reactions on a basis different from that currently held. The reversibility of the precipitin reaction, in the sense that a precipitate may be shifted from one region of the reaction range to another by addition of either component, suggested a similar possibility for the closely related agglutinin reaction. It was predicted that if, for example, Type I pneumococci were agglutinated with a large excess of antibody, and that if the

TABLE 6*
Molecular composition of specific precipitates from rabbit antisera

| ANTIGEN | EMPIRICAL COMPOSITION OF SPECIFIC PRECIPITATE | | | | COMPOSITION OF SOLUBLE COMPOUNDS IN INHIBITION ZONE |
|------------------------------------------------------------|-----------------------------------------------|----------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------|-----------------------------------------------------------------|
| | At extreme antibody excess | At antibody excess end of equiv- alence zone | At antigen excess end of equiv- alence zone | In inhibition zone | |
| Crystalline egg albumin ⁽¹⁾ | EaA ₅ | EaA ₅ | Ea ₂ A ₅ → EaA ₂ | EaA ₂ → (EaA) | |
| Dye egg albumin ⁽²⁾ | (DEaA ₅) | (DEaA ₅) | DEa ₂ A ₅ → DEa ₄ A ₂ | DEa ₄ A ₂ → DEa ₂ A (?) | |
| Crystalline serum albu- min ⁽³⁾ | SaA ₆ | SaA ₄ | SaA ₃ → SaA ₂ | SaA ₂ → (SaA) | |
| Thyroglobulin ⁽⁴⁾ | TgA ₁₀ | TgA ₁₄ | TgA ₁₀ → TgA ₂ | TgA ₂ → (TgA) | |
| Type III pneumococcus ⁽⁵⁾ | SA | S ₂ A ₂ | S ₂ A → S ₄ A | S ₄ A → S ₂ A | |

* Reprinted from J. Am. Chem. Soc. 60, 242 (1938). A = antibody; S = minimum polysaccharide chain weight reacting. Data in parentheses are somewhat uncertain.

(1) Heidelberger, M., and Kendall, F. E.: J. Exptl. Med. 62, 697 (1935).

(2) Heidelberger, M., and Kendall, F. E.: J. Exptl. Med. 62, 487 (1935).

(3) Kabat, E. A., and Heidelberger, M.: J. Exptl. Med. 66, 229 (1937).

(4) Stokinger, H. E., and Heidelberger, M.: J. Exptl. Med. 66, 251 (1937).

(5) Heidelberger, M., and Kendall, F. E.: J. Exptl. Med. 65, 647 (1937)

excess of antibody were removed by thorough washing and the agglutinated pneumococci were resuspended in physiological salt solution, they would reagglutinate immediately into larger clumps on addition of appropriate amounts of fresh Type I pneumococci or of Type I specific polysaccharide. It was also predicted that no change would occur on addition of, for example, Type II pneumococci or Type II specific polysaccharide, although conditions of electrical potential and cohesive force (34, 37) would be nearly identical. These predictions were fully verified (14). The experiments were considered to indicate that in this way the process of specific bacterial agglutination had been, in a sense, interrupted at an

early stage (much antibody and few pneumococci), and that resumption of the agglutination process under controlled conditions showed that aggregation into large clumps was merely a continuation of the chemical union of multivalent antigen on the effective bacterial surface with multivalent antibody on other effective bacterial surfaces. The assumption of an initial static antigen-antibody union and the subsequent building up of large aggregates through a combination of cohesive force and salt-lowered potential are thus contrary to actual experimental data. These data have been criticized (29), but the alternative explanation that dissociation of antibody could have occurred was eliminated by washing the initially agglutinated pneumococci until the supernatant no longer agglutinated added pneumococci, and also by the obviously greatly larger clumps formed in the reagglutination.

Not only, then, is there evidence that the process of specific bacterial agglutination (and hence the precipitin reaction) differs from other aggregating systems long used as analogies, in that it behaves as a single chemical reaction from beginning to end (see also Topley, Wilson, and Duncan (42)), but the same experiments place the function of salts in another light. The effect of salts may be interpreted as the purely secondary one of providing ions for the ionized salt complexes in which form antibody probably reacts, and, in addition, of minimizing electrostatic effects due to the presence of many ionized groupings on the particles, effects which might interfere with the completion of particulation by chemical interaction.

Whether or not the initial bimolecular antigen-antibody reaction on the bacterial surface can take place in the absence of electrolyte, the reactants carry ionized groups, and it is evident that the succeeding competing bimolecular interactions between polysaccharide molecules on partly sensitized cells and additional antibody in solution or on other cells would soon result in the formation, in the absence of electrolyte, of particles carrying large numbers of ionized groups. Coulomb forces on such particles are known to cause abnormally great viscosities and Donnan effects, so that it would not be surprising if these forces would prevent the continuation of the chemical reactions resulting in the completion of what is commonly recognized as specific bacterial agglutination. Only when the effect of these forces is eliminated by a sufficient ionic atmosphere, on addition of electrolyte, is it possible to obtain significant figures for viscosity, osmotic pressure, sedimentation constants, and the like. To ascribe a similar rôle to electrolytes in specific bacterial agglutination would seem reasonable and consistent, for after reduction of the Coulomb forces the growing particles could again interact chemically, and the process of agglutination be completed. An analogous explanation of the

function of salts in the reaction between the heterogenetic antigen of beef kidney and antibody had already been given by Brunius (2), but this was not known until after the publication of our paper (14). Possibly these immunochemical studies will contribute to an understanding of the mechanism of the agglutination of suspended particles in general.

Enough, then, has been cited to show the utility of the quantitative theories of the precipitin and agglutinin reactions as working hypotheses and temporary expedients until additional precise information makes a better structure possible. A beginning has been made, but only through the continued coöperative efforts of the organic chemist, the physical chemist, and the biologist is it likely that there will result a final understanding of the complex chemical phenomena underlying immunity.

REFERENCES

- (1) AVERY, O. T.: *Naturwissenschaften* **21**, 777 (1933).
- (2) BRUNIUS, F. E.: *Chemical Studies on the True Forssman Hapten, the Corresponding Antibody, and their Interaction*. Dissertation, Stockholm, 1936.
- (3) CULBERTSON, J. T.: *J. Immunol.* **28**, 279 (1935).
- (4) FELTON, L. D.: *J. Infectious Diseases* **43**, 543 (1928), and other papers.
- (5) FELTON, L. D.: *J. Immunol.* **22**, 453 (1932).
- (6) FELTON, L. D., AND KAUFFMANN, G.: *J. Immunol.* **25**, 165 (1933).
- (7) GOODNER, K., AND HORSFALL, F. L., JR.: *J. Exptl. Med.* **66**, 437 (1937).
- (8) GOODNER, K., HORSFALL, F. L., JR., AND BAUER, J. H.: *Proc. Soc. Exptl. Biol. Med.* **34**, 617 (1936).
- (9) HEIDELBERGER, M.: *Physiol. Rev.* **7**, 107 (1927); *Chem. Rev.* **3**, 403 (1926-27).
- (10) HEIDELBERGER, M.: Chapter 14 in B. Harrow and C. P. Sherwin's *Textbook of Biochemistry*, Saunders, Philadelphia (1935); *Bact. Rev.*, in press.
- (11) HEIDELBERGER, M.: *J. Am. Chem. Soc.* **60**, 242 (1938).
- (12) HEIDELBERGER, M., AND GOEBEL, W. F.: *J. Biol. Chem.* **70**, 613 (1926); **74**, 613 (1927).
- (13) HEIDELBERGER, M., AND KABAT, E. A.: *J. Exptl. Med.* **60**, 643 (1934); **67**, 545 (1938).
- (14) HEIDELBERGER, M., AND KABAT, E. A.: *J. Exptl. Med.* **65**, 885 (1937).
- (15) HEIDELBERGER, M., AND KABAT, E. A.: *J. Exptl. Med.* **67**, 181 (1938).
- (16) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Exptl. Med.* **50**, 809 (1929).
- (17) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Biol. Chem.* **95**, 127 (1932).
- (18) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Exptl. Med.* **57**, 373 (1933).
- (19) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Exptl. Med.* **61**, 559 (1935).
- (20) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Exptl. Med.* **61**, 563 (1935).
- (21) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Exptl. Med.* **62**, 467 (1935).
- (22) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Exptl. Med.* **62**, 697 (1935).
- (23) HEIDELBERGER, M., AND KENDALL, F. E.: *J. Exptl. Med.* **64**, 161 (1936).
- (24) HEIDELBERGER, M., KENDALL, F. E., AND SCHERP, H. W.: *J. Exptl. Med.* **64**, 559 (1936).
- (25) HEIDELBERGER, M., KENDALL, F. E., AND THEORELL, T.: *J. Exptl. Med.* **63**, 819 (1936).

- (26) HEIDELBERGER, M., AND PEDERSEN, K. O.: J. Exptl. Med. **65**, 393 (1937).
- (27) HEIDELBERGER, M., PEDERSEN, K. O., AND TISELIUS, A.: Nature **138**, 165 (1936).
- (28) HOOKER, S. B., AND BOYD, W. C.: J. Gen. Physiol. **17**, 341 (1934); Proc. Soc. Exptl. Biol. Med. **32**, 1104 (1935) and later papers.
- (29) HOOKER, S. B., AND BOYD, W. C.: J. Immunol. **33**, 337 (1937).
- (30) KABAT, E. A., AND HEIDELBERGER, M.: J. Exptl. Med. **68**, 229 (1937).
- (31) KABAT, E. A., AND PEDERSEN, K. O.: Science **87**, 372 (1938); KABAT, E. A.: J. Exptl. Med. **69**, 103 (1939).
- (32) cf. also MARRACK, J. R.: Chemistry of Antigens and Antibodies. His Majesty's Stationery Office, London (1934; second edition, 1938).
- (33) MARRACK, J. R., AND CARPENTER, B. R.: Brit. J. Exptl. Path. **19**, 53 (1938).
- (34) NORTHRUP, J. H., AND DE KRUIF, P.: J. Gen. Physiol. **4**, 629, 639, 655 (1922).
- (35) PAPPENHEIMER, A. M., JR., AND ROBINSON, E. S.: J. Immunol. **32**, 291 (1937).
- (36) PENNELL, R. B., AND HUDDLESON, L. F.: J. Exptl. Med. **68**, 73, 83 (1939).
- (37) SHIBLEY, G. S.: J. Exptl. Med. **40**, 853 (1924).
- (38) STOKINGER, H. E., AND HEIDELBERGER, M.: J. Exptl. Med. **68**, 251 (1937).
- (39) TAYLOR, G. L., ADAIR, G. S., AND ADAIR, M. E.: J. Hyg. **34**, 118 (1934).
- (40) TISELIUS, A.: J. Exptl. Med. **65**, 641 (1937).
- (41) TISELIUS, A., AND KABAT, E. A.: Science **87**, 416 (1938); J. Exptl. Med. **69**, 119 (1939).
- (42) TOPLEY, W. W. C., WILSON, J., AND DUNCAN, J. T.: Brit. J. Exptl. Path. **16**, 116 (1935).
- (43) For further details and references cf. HEIDELBERGER, M., TURNER, J. C., AND Soo Hoo, C. M.: Proc. Soc. Exptl. Biol. Med. **37**, 734 (1938).
- (44) WELLS, H. G.: Chemical Aspects of Immunity, 2nd edition, p. 139. The Chemical Catalog Co., Inc., New York (1929).

THE POLAR GROUPS OF PROTEIN AND AMINO ACID SURFACES IN LIQUIDS¹

H. A. ABRAMSON

*Department of Physiology, College of Physicians and Surgeons, Columbia University,
New York City; The Medical Service of Dr. George Baehr and the Laboratories of
the Mount Sinai Hospital, New York City; The Biological Laboratory, Cold Spring
Harbor, Long Island, New York*

M. H. GORIN

*The Medical Service of Dr. George Baehr and the Laboratories of the Mount Sinai
Hospital, New York City; The Biological Laboratory, Cold Spring Harbor,
Long Island, New York*

AND

L. S. MOYER

*Department of Botany, University of Minnesota, Minneapolis, Minnesota; The
Biological Laboratory, Cold Spring Harbor, Long Island, New York*

Received January 30, 1939

The most striking characteristic of the polar groups of protein molecules and of protein surfaces or films is the presence of comparatively large numbers of amino and carboxyl radicals. While many groups besides the COO^- and NH_3^+ groups serve to determine the general polarity of the molecule, these and other acidic and basic groups alone primarily determine the net charge of the protein under ordinary conditions.

A most sensitive indicator which gives special information regarding the number and orientation of the COO^- and NH_3^+ groups of protein surfaces is the electric mobility. Thus, it will be shown that, for the cases so far investigated, when

$$\sum \text{COO}^- = \sum \text{NH}_3^+$$

on each molecule over a time average, the electric mobility is zero, and the protein or the protein surface is called isoelectric. Near the isoelectric point the slope of the electric mobility-pH curve is greater than at regions far removed from the isoelectric point. An exception to this rule is found

¹ Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

only in fairly alkaline solutions where the possibility of the instability of protein exists. On the other hand, in acid solutions not only does the difficulty of interpretation of hydrogen-ion activity measurements exist, but also the electric mobility curve of proteins flattens out so that the change in electric mobility, v , with pH becomes quite unimportant. In view of the fact that at or near the isoelectric point the value of dv/dpH is usually a maximum, experiments designed to investigate changes in the

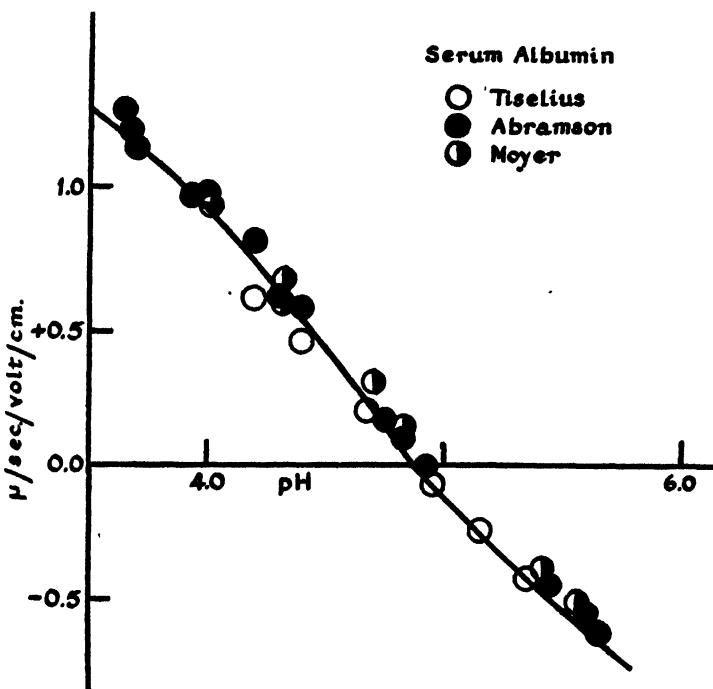


FIG. 1. The electric mobility of horse serum albumin as a function of pH, in 0.02 M acetate buffers at 20°C. There is no significant difference between the data for adsorbed (Abramson, Moyer) and dissolved (Tiselius) protein. The smooth curve was drawn free-hand through the points.

number and distribution of the polar groups of proteins and other organic amphotytes incidental to changes in state have always included measurements made at or near the isoelectric point. In this communication, systems will be described which summarize the changes in number and orientation of the amphoteric polar groups of proteins after adsorption at solid-liquid and liquid-liquid interfaces. In addition, data will be presented which will indicate the nature of the changes in the isoelectric point

occurring at the surfaces of organic ampholytes like amino acids incidental to crystallization.

On the basis of recent data, the changes in the polar groups of organic ampholytes incidental to surface formation may be classified into three groups:

Group I: This group deals primarily with proteins after adsorption by quartz and other particles. Within the limits of the method employed, little or no change occurs in the number and orientation of the acid and basic groups, as indicated by the constancy of the electric mobility and the isoelectric point.

Group II: This group, as group I, is composed of proteins. In group II slight changes in the number and orientation of polar groups occur, as evidenced by a change in the isoelectric point of the surface and the mobility-pH curves.

Group III: Here profound changes in the nature of the orientation and in the number of polar groups occur with surface formation. The best examples are the surfaces of amino acid crystals.

The first part of this paper will be concerned with these three groups. In the second part the relationship between titration curves and electric mobility of proteins will be discussed. Finally, the calculation of an "effective" radius for protein molecules from mobility data and its interpretation will be presented.

I. FILM FORMATION BY AMPHOLYTES

Group I

Although numerous experiments on adsorbed protein surfaces or particles of denatured protein had previously been performed by Pauli and Flecker (38), Walpole (49), Brossa and Freundlich (15), Loeb (22), Hitchcock (21), Reinders and Bendien (40), Freundlich and Abramson (19), Prideaux and Howitt (39), and Mattson (23), a direct comparison of the electric mobilities of dissolved proteins and of protein surfaces was not possible until Tiselius published his first group of moving boundary data in 1930.

The first direct comparison of the electric mobility of a dissolved protein with that of an adsorbed protein under precisely similar conditions was made by Abramson in 1932 (5). In figure 1 are given the data of Tiselius for the dissolved form of horse serum albumin and those of Abramson (5) and Moyer (29) for quartz and collodion particles, respectively, coated with this protein. All measurements presented were made in 0.02 *M* acetate buffers by methods described by Abramson (8), Abramson and Moyer (11), Abramson, Moyer and Voet (13), and Moyer (24, 28).

It may be seen from figure 1 that the points fall together on a smooth curve within the limits of experimental error, and therefore that the electric mobilities of dissolved serum albumin (in 0.02 *M* acetate buffer)

are the same as the electric mobilities of quartz and collodion particles covered with complete films of serum albumin. It is not known whether these films are monomolecular or polymolecular. The experiments were performed with a sufficient concentration of protein to yield at least complete films. Neither the size nor the nature of these particles covered with films of protein influenced the electric mobility of the protein-covered particles. Indeed, it was found in acetate buffers and in solutions containing other electrolytes that the mobility of a microscopically visible, protein-covered quartz particle having a diameter of 1 micron was exactly equal to the electric mobility of a particle similarly covered with protein but 10, 100, or even 1000 times larger (4, 8, 10). More striking, the ratio of the electrophoretic mobility of the particle to the electroosmotic mobility of the liquid along a flat surface covered with the same protein is equal to 1.0 (4, 6, 8, 16, 17, 27, 30, 33). In other words, protein films of this kind had mobilities independent of the size, shape, and orientation of the particle. All of these results may be visualized as follows: A single dissolved serum albumin molecule migrating in the same field with a quartz particle 1 micron in diameter, covered with a film of serum albumin, and a baseball covered with a film of this protein in the same electric field would migrate at the same rate. The physical reality of the phenomenon is approached by imagining that, rather than the particle itself in motion, we have the water streaming past the particle, with the result that the water has the same speed whether streaming past a single albumin molecule or passing along a wall coated with serum albumin.

The electric mobility, v , of a sphere or of other highly simplified models of charge density σ and radius r may, under certain circumstances, be represented by an expression of the form,

$$v = f(\sigma, r)$$

in a given medium (8). In order for the electric mobility to remain the same after adsorption of the protein on a surface, both the charge density and the radius must change in such a way that they exactly compensate one another, or else both must remain the same after adsorption as in the freely dissolved state. This maintenance of constant electric mobility for a single dissolved molecule and the adsorbed film could hardly be achieved by such a compensation mechanism. It is necessary, therefore, to conclude that in this instance there was no important change in the orientation or activity of the carboxyl and amino groups incidental to the adsorption of the protein and the formation of the complete film. Furthermore, a large degree of distortion in the molecules of the outermost layer could not have occurred. That is, the effective radii of curvature of the adsorbed protein molecules are essentially the same as the effective radii of curvature

of the dissolved molecules. The implications of this conclusion have been discussed by Abramson (5) and by Moyer (27, 28, 31).

An investigation by Gorin, on the other hand, has shown that with the charge remaining constant a simple elongation of the molecule, which would amount to a distortion, could occur without changing the mobility very much. However, it is unlikely that the net charge would remain constant upon distorting the molecule to any appreciable extent. This is supported both by the "salt-bridge" theory and by the work on egg albumin to be presented later in this paper.

Another protein which has since been investigated is pseudoglobulin. A comparison of the data on dissolved pseudoglobulin found by Tiselius

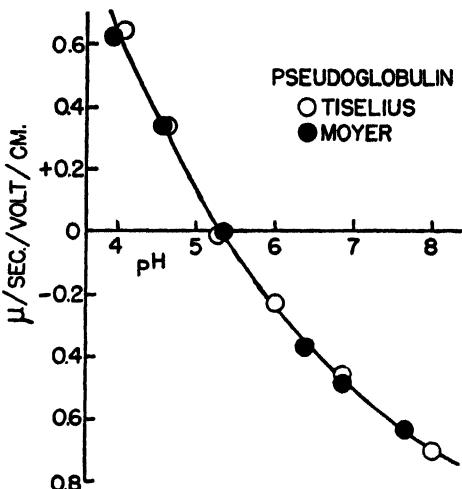


FIG. 2. The electric mobility curve of horse serum pseudoglobulin in buffers of a constant ionic strength of 0.1 at 20°C. Notice that a single curve fits the data for dissolved protein (Tiselius) as well as for adsorbed protein (Moyer).

(47) with the moving boundary method with those for adsorbed pseudoglobulin films on collodion particles by Moyer (29) reveals the same remarkable constancy of electric mobility from pH 4 to pH 8. This comparison is made in figure 2. This protein and serum albumin appear to be the only ones for which data are available for exact comparisons of the dissolved and adsorbed forms. Other data, however, give indirect evidence that similar cases may exist.²

² The data on the mobility of particles coated with gelatin and its relation to the titration curves of dissolved gelatin furnish indirect evidence that gelatin is another protein that displays this remarkable property. They will be presented in the section on titration curves and mobility.

Further experiments may be directed toward studying films of the various globulin fractions now made available by the more recent experiments of Tiselius (48) with the Toepler method of analysis of the moving boundary. The comparison of the films of these proteins, and of their mixtures, with dissolved molecules will be especially interesting, not only from the immunological point of view but also for a further analysis of the surface composition of living cells. In the case of systems where the protein itself can be traced by the behavior of protein-coated quartz particles, further elucidation of the problem of mutual adsorption of protein, or of selective adsorption of protein, may be sought. Experiments in this field have been begun by Moyer and Moyer (35). Earlier investigations of Abramson (2) on the behavior of quartz and paraffin oil droplets in whole serum are a case in point. Other biochemical and biological applications have been recently discussed by Abramson (8), Abramson and Moyer (11), and Moyer (25, 26, 31).

Group II

In the case of egg albumin the electrophoretic behavior of the adsorbed film differs slightly, but definitely, from that of the molecules in solution. The results of Moyer (28), Abramson (5), and Tiselius (46) are shown in figure 3. The solid line, labeled "adsorbed" in the figure, represents the smoothed data of Moyer for films of egg albumin on quartz, glass, oil droplets, collodion, and carbon. Earlier work of Abramson for films on quartz is represented by the closed circles. Data of Tiselius for dissolved egg albumin are shown by the open circles.

It may be observed that the smoothed curve for adsorbed egg albumin, found by Moyer and with which we agree, is essentially parallel to that of Tiselius over the range of pH investigated. However, it is definitely shifted to the right on the pH scale; the isoelectric point is at about pH 4.82 instead of pH 4.55. In the same figure is shown the smoothed curve by Moyer (broken line) for the mobilities of particles of surface-denatured egg albumin prepared by shaking. This curve lies still further to the right (isoelectric point at about pH 5.0) than that of the adsorbed normal egg albumin. The discrepancy between the points of Abramson and the smooth curve of Moyer is probably due to the fact that Abramson used quartz particles which had been in long contact with water. This is borne out by the fact that Moyer's curve for egg albumin films on silica gel particles is also anomalous and within the limits of error of Abramson's points.

It is evident that the electrokinetic properties of air-denatured egg albumin are considerably different from those of molecules of egg albumin in solution. The differences in the isoelectric point are to be correlated

with changes in the relative strengths of the acidic and basic groups incidental to denaturation. Alterations in the shape of the molecule may also occur. These would be expected to result in a curve of different shape, but would only indirectly affect the isoelectric point (28). No definite conclusions may be drawn however, for changes in the shape of the curve must also accompany changes in the number and relative strengths of the available acidic and basic groups.

The "salt-bridge" theory of Stearn and Eyring (43) is useful in considering changes in the relative strengths of acidic and basic groups that

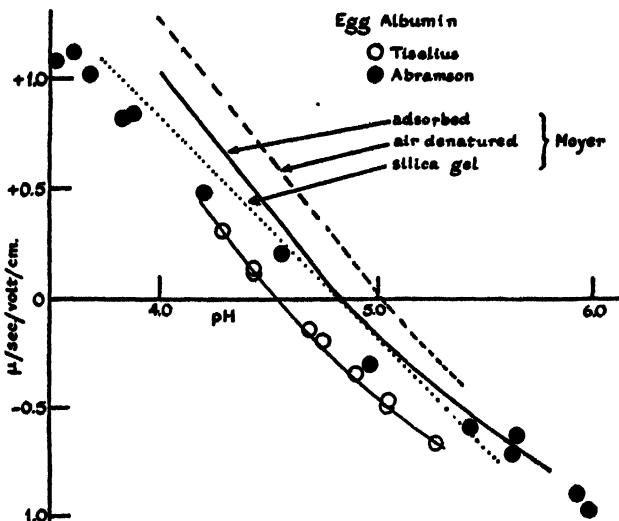


FIG. 3. Comparison of the data of Tiselius for dissolved egg albumin with the results of Abramson and of Moyer for adsorbed films of this protein. The smooth curve of Moyer represents the best curve drawn through data obtained with five preparations of egg albumin adsorbed on five different surfaces. For comparison, the smoothed curves for air-denatured egg albumin and egg albumin adsorbed on silica gel have been included. All data were obtained at $\mu = 0.02$ and $T = 20^\circ\text{C}$.

might occur upon denaturation. According to these authors, bonds between NH_3^+ and COO^- on opposite side chains are at least partially responsible for the maintenance of the folded state of the native protein. Upon denaturation, according to the widely accepted theories of Wu (51), Astbury (4), and others, the molecule unfolds and therefore these salt bridges are broken, say Stearn and Eyring. It is to be expected then, from this picture, that shifts in the relative strengths of the amino and carboxyl groups would occur upon denaturation, for important changes would take place in the position of some of these groups in the molecule.

Since the isoelectric point of the adsorbed films lies between that of the dissolved and denatured egg albumin, the molecules in these films may be considered to be "partially denatured." It is not to be inferred without further evidence that all the physical, chemical, and biological properties of the molecules in the adsorbed film lie between those of the dissolved and surface-denatured form. Rather, the term "partially denatured" is used to indicate that the difference between dissolved and adsorbed egg albumin may be due to the instability of this native protein toward surface denaturation.

Some difficulty may be experienced in classifying the electrokinetic behavior of insulin crystals. We have found no data in the literature for the electric mobility of insulin in the dissolved state, but data are

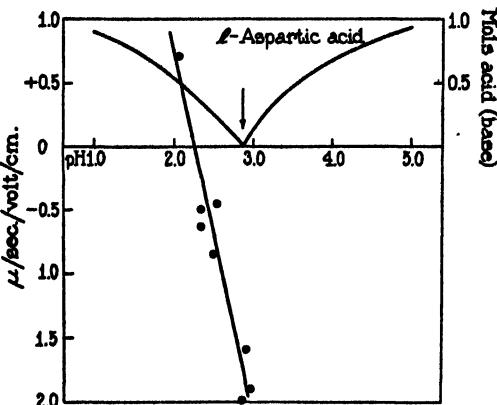


FIG. 4. A comparison of the electric mobilities of aspartic acid crystals in hydrochloric acid solutions and the titration curve of dissolved aspartic acid. The arrow indicates the position of the isoelectric point of dissolved aspartic acid. ●, hydrochloric acid. Smooth curve, titration.

available for the electrophoretic behavior of adsorbed insulin and amorphous particles of insulin suspended in buffers (50). If amorphous insulin, that is, insulin precipitated rapidly from solution at the isoelectric point, and films of insulin on quartz are compared directly in the same solution near the isoelectric point, the mobilities are, within the limits of error, identical. Crystals of insulin, however, studied under the same conditions, show a divergence from this behavior. The isoelectric point of one preparation of insulin crystals studied by Wintersteiner and Abramson differed from that of the adsorbed film of insulin by approximately 0.35 pH unit. The general shape of the mobility-pH curve of the insulin crystals is that of a typical amphotyte. It is not known in this instance

how much adsorbed insulin was present on the surface of the crystalline insulin, because some dissolved insulin is always present even at or near the isoelectric point. We may perhaps tentatively include insulin in the second group, because the difference between adsorbed and crystalline insulin is not marked. This difference might have been greater if no dissolved insulin at all had been present. On the basis of their state of aggregation, a better classification would perhaps include insulin crystals in the group of amphoteric crystals, the amino acids, which are now to be discussed.³

Group III

The recent work of Abramson and Moyer (12) on the electrokinetic properties of amino acid crystals furnishes data for a third type of behavior,—the very different isoelectric points and mobilities of dissolved and crystalline amino acids. It is offered, not because it is exceptional from a general point of view, but because it allows, by contrast, a better insight into the unique behavior of the protein surfaces included in the first group.

Crystals of aspartic acid, cystine, and tyrosine were suspended in hydrochloric acid, and, in the case of the last two, also in buffers at two ionic strengths. Their mobilities were determined in the flat microelectrophoresis instrument of Abramson (1, 3, 8, 9) over a wide range of pH. Results are given in figures 4 to 6 and are compared with those predicted for the dissolved molecules either from titration curves or solubility determinations. It is seen from these figures that the difference between the isoelectric points of dissolved and crystalline amino acids ranges from 0.5 pH unit in the case of aspartic acid to more than 3.0 units in the case of tyrosine, and that the isoelectric points of the crystals of the three substances are very close together (about pH 2.4), while those of the corresponding dissolved amino acids are far apart. The curves obtained for the crystals show no inflections at the isoelectric points of both the dissolved and the crystalline forms.

The effects of salts on the mobilities of cystine and tyrosine were investigated. It may be seen from figures 4 to 6 that increasing the ionic strength decreased the mobility at a given pH value. No data of this

³ One of us (H. A. A.) has measured the electric mobility of hemoglobin crystals at the isoelectric point of dissolved hemoglobin in a saturated solution of hemoglobin in phosphate buffers. Under these conditions the hemoglobin crystals were negatively charged. An insufficient number of experiments and the presence of impurities, even though the preparation had been recrystallized several times, prevent further comments from being made at this time, but the data illustrate that protein crystals, even in the presence of concentrated, saturated solutions of the protein, may not have the electrokinetic properties of the dissolved molecules.

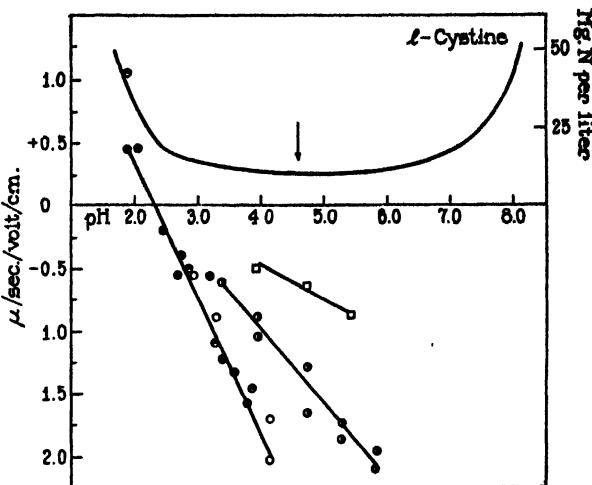


FIG. 5. The electric mobilities of cystine crystals suspended in hydrochloric acid and in sodium acetate buffers. The arrow points to the isoelectric point of dissolved cystine. ●, ○, hydrochloric acid; ⊖, N/200 sodium acetate; □, N/200 sodium acetate + N/5 sodium chloride. Smooth curve, solubility (Sano).

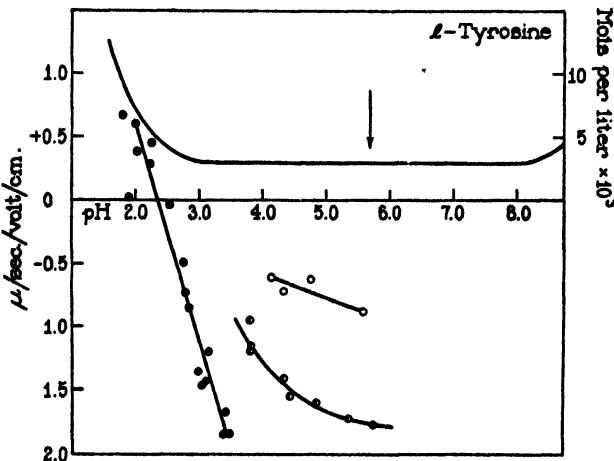


FIG. 6. The electric mobilities of tyrosine crystals suspended in hydrochloric acid and in sodium acetate buffers. The arrow points to the isoelectric point of dissolved tyrosine. ●, hydrochloric acid; ⊖, N/200 sodium acetate; ○, N/200 sodium acetate + N/8 sodium chloride. Smooth curve, solubility (Hitchcock).

type were obtained on the positive side of the isoelectric point of the crystals, nor are there any data very close to the isoelectric point. Available measurements appear to extrapolate to the value of the isoelectric

point obtained for hydrochloric acid solutions. However, the possibility of the existence of bends in the curves at lower values of pH, with corresponding shifts in the isoelectric points, must be taken into consideration.

The tremendous shift in the isoelectric points of the surface of amino acid crystals is probably not to be explained merely by a corresponding change in the relative strengths of the acidic and basic groups. That

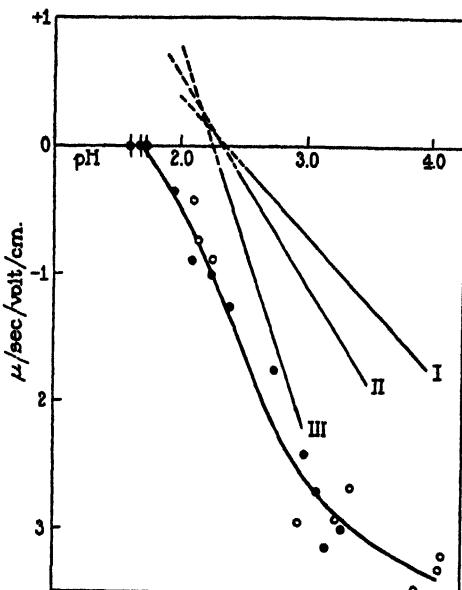


FIG. 7

FIG. 7. A comparison of the electric mobilities of emulsion droplets of simple alkylbenzenes with those of the surfaces of crystalline amino acids. ○, *n*-propylbenzene; ●, ethylbenzene; I, cystine (crystal); II, tyrosine (crystal); III, aspartic acid (crystal).

FIG. 8. This figure illustrates the influence of polyvalent ions on the isoelectric point. By dividing the ordinate values by 13 the approximate electrical mobility may be obtained. The abscissa represents molarity of (1) sodium ferrocyanide, (2) sodium chloride, and (3) lanthanum chloride. Note that even at very low concentrations of polyvalent ions, the effect is much more than was observed at much higher concentrations of uni-univalent salts.

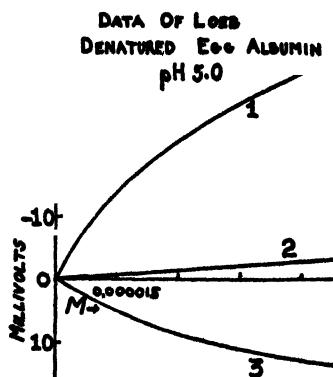


FIG. 8

large changes in the relative strengths of the acidic and basic groups of the crystal surfaces, compared with those of the dissolved molecules, might occur is not to be denied. It is unlikely, however, that the charge at the surface of the amino acid crystals is solely determined by the amphoteric properties (ionogenic) of the surface. Negative ions may also be adsorbed at the surface. The net charge would then result from

a competition of hydrogen ions, or positively charged amino acid molecules, and the anions in the solutions, with a shift of the isoelectric point to lower pH values.

In the same paper Abramson and Moyer (12) reported investigations of the mobilities of droplets of *n*-propylbenzene and ethylbenzene in solutions of hydrochloric acid (pH 2 to 4). Their results are shown in figure 7 (taken from figure 6 of Abramson and Moyer) along with the corresponding results for cystine, tyrosine, and aspartic acid crystals already discussed. The non-amphoteric surfaces of *n*-propylbenzene and ethylbenzene show some similarities to those of the amino acid crystals. They differ primarily in the apparent absence of a reversal of charge at lower values of the pH. The mobilities are, in addition, greater at corresponding values of the pH.

II. TITRATION CURVES AND MOBILITY

The demonstration (5) that a direct proportionality exists between titration curves and electric mobility in simple systems furnished a basis for further applications of electrokinetic data to the investigation of the nature of the surfaces of protein molecules and of adsorbed films of proteins. For if, at constant ionic strength in simple systems, the average charge of the molecule is known to be almost completely determined by the average number of equivalents of hydrogen or hydroxide ions bound, it follows that changes in the isoelectric point incidental to surface film formation must be directly correlated with changes in the relative strengths of the amino and carboxyl groups. Furthermore, the application of modern theories of electrolytes to the interpretation of the entire mobility curve would be almost impossible if the average charge on the molecule were not known.

It must be emphasized that the preceding remarks are made specifically for protein solutions containing only simple electrolytes which do not form protein complexes (5, 7, 8). Loeb (22), for example, showed that particles of casein, denatured egg albumin, and gelatin were markedly influenced at the isoelectric point by polyvalent ions. Thus, a trace of sodium ferrocyanide ($M/65,000$) increased the mobility of isoelectric egg albumin particles to values unusually high for proteins at the pH investigated. Other data may be found among the investigations of Tiselius (46), who found that Ba^{++} shifted considerably the isoelectric point of dissolved egg albumin. The effect of citrate ion on the isoelectric point is well exemplified by the data of Ottenberg and Stenbeck (37), who found that the isoelectric point of typhoid agglutinin in solution was shifted about 1 pH unit by citrate buffers. The effects of other ions have been investigated by Smith (41, 42) with particles coated with egg albumin.

While data obtained on the electric mobility of proteins in the presence of polyvalent ions may be of significance in evaluating specific ion effects, considerable caution should be exercised before interpreting mobility data obtained in this way. At present it seems most pertinent for further progress to investigate the electric mobility and charging process of proteins in the presence of salts which do not shift the isoelectric point markedly. It might be of interest to investigate proteins in the presence

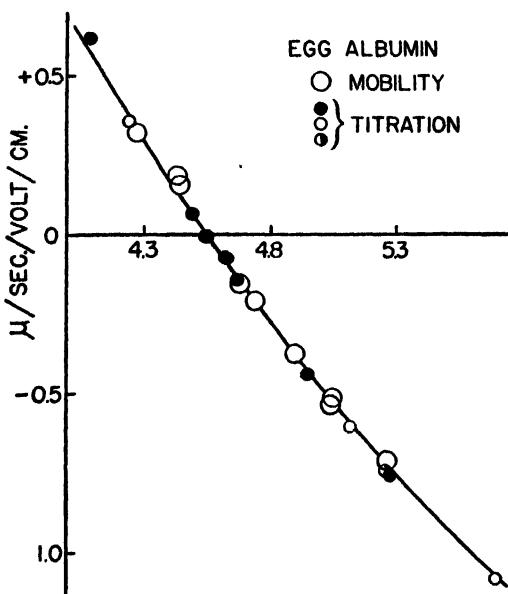


FIG. 9. Comparison of the titration data of Moyer and Abels (32) with the mobility data of Tiselius (48) for dissolved egg albumin at 25°C. The different kinds of circles represent different preparations of egg albumin. All determinations were made at a constant ionic strength of 0.02.

of perchlorate ion, because it, of all anions, has shown the least tendency to form complexes in other systems.

To demonstrate the proportionality between the combining power of proteins with acids and bases and their electric mobilities, various methods of plotting have been employed. The method employed here is that used by Abramson (5). In a plot with the number of equivalents of acid (base) bound per gram of protein as ordinate and the pH as abscissa, the zero point on the ordinate is found by shifting the coördinates vertically until the curve goes through zero at the isoelectric point as determined by electrophoresis. The mobility data are then compared with the titra-

tion data by taking any single point on the smoothed mobility curve and multiplying the mobility by a factor that will make the two curves correspond. If, when all the experimental points are multiplied by this same factor, they fall on the titration curve within the limits of error, this proportionality is to be considered as demonstrated. The most carefully worked out case is that of Moyer and Abels (32) for egg albumin. It is presented in figure 9, which was constructed as described above. In this instance, both sets of data had been measured at the same, constant ionic strength.

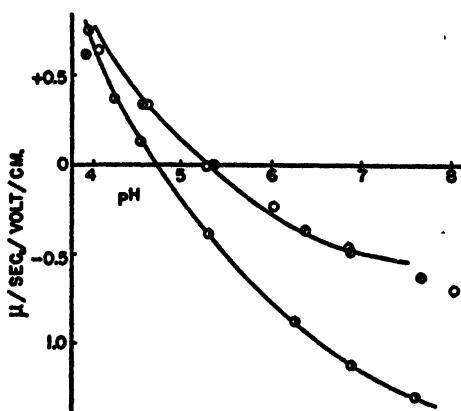


FIG. 10. Electric mobility curves of serum albumin (●) and pseudoglobulin (dissolved, ○; adsorbed, ◎) in phosphate and acetate buffers at $\mu = 0.1$. The smooth curves are the titration curves for these proteins drawn to fit the mobility data. The scales for the titration curves are different. The ordinate gives equivalents of hydrogen ion bound when multiplied by 2×10^{-4} for pseudoglobulin and by 3×10^{-4} for serum albumin. This shows that over a wide pH range the electric mobility of these proteins is proportional to their combining power with hydrogen (or hydroxide) ions. (From Moyer and Abramson (34)).

There has been considerable confusion about the definition of the isoelectric point for ampholytes. Physically, this point refers to conditions under which the mobility of the ampholyte is zero, and therefore it can only be strictly defined in terms of the charge on the ampholyte and determined by electrokinetic or conductance measurements (8, 12). Thermodynamic definitions of the isoelectric point involve assumptions which may or may not be strictly valid but certainly cannot be taken for granted until demonstrated in each individual case. For proteins in simple systems containing only uni-univalent electrolytes, the assumptions are very nearly valid in cases so far investigated, and the thermodynamic definition may be considered applicable. It seems desirable, however, to define

the isoelectric point more generally, as suggested by one of us (8). According to this point of view, the isoelectric point of an ampholyte, whether dissolved or constituting a surface film, may be defined as a reference concentration (H^+ , for example) at which

$$\frac{1}{T} \int_0^T dt \left\{ \sum_i n_i z_i(e) + \sum_i n_i z_i(-e) \right\} = 0 \quad (1)$$

$(T \gg \tau)$

Here e is the electronic charge and n_i is the number of ions of the i^{th} type, of valence z_i , at the surface during the time dt . T represents the time of observation and τ the life period of an ion at the surface. The conditions relating to the life period, τ , of the ion may be fulfilled by observing a large number of fluctuations during the period of measurement, i.e., a statistical sample.

Other cases in which the proportionality has been demonstrated between acid (base) bound and mobility are those of serum albumin and pseudoglobulin (34). They are shown in figure 10 and drawn as described above. Other cases referred to at the beginning of the section are those of gelatin and deaminized gelatin. There are no data available for the mobility of the dissolved forms, but it has been shown that a simple proportionality exists between the mobility of quartz particles coated with gelatin and the acid(or base)-binding of the dissolved form. It was thus shown indirectly that gelatin is another protein for which the mobilities of the dissolved and adsorbed forms must be very nearly identical.

III. INTERPRETATION OF ELECTRIC MOBILITIES

Calculation of protein radii

The potential at the surface of a particle of known net charge Q in solutions of electrolytes depends upon the size and shape of the particle, the charge distribution over its surface, and the nature of the electrical double layer surrounding the particle. For a rigid sphere, Debye and Hückel (18) derived the following expression for the surface potential, ζ ,

$$\zeta = \frac{Q}{Dr(\kappa r + 1)} \quad (2)$$

where D is the dielectric constant of the medium, r the radius of the molecule, Q the charge on the molecule, and $\kappa = 0.328 \times 10^6 \sqrt{C}$ for uni-univalent electrolytes at 25°C. The mobility, v , is related to ζ as follows,

$$v = \frac{Drf(\kappa r)}{6\pi\eta} \quad (3)$$

where η is the viscosity of the medium (8, 27), and, according to Henry (20),⁴

$$f(\kappa r) = \left\{ 1 + \frac{(\kappa r)^2}{16} - \frac{5(\kappa r)^4}{48} - \frac{(\kappa r)^6}{96} + \frac{(\kappa r)^8}{96} - \left[\frac{12(\kappa r)^4}{96} - \frac{(\kappa r)^6}{96} \right] \int_{\infty}^{\kappa r} \frac{e^{-t}}{t} dt \right\}$$

The function, $f(\kappa r)$, varies between 1 and 1.5 as κr goes from zero to infinity. From equations 2 and 3 it follows that

$$v = \frac{Qf(\kappa r)}{6\pi\eta(\kappa r + 1)} \quad (4)$$

The assumptions involved in the Debye-Hückel development are: (1) Charge is uniformly distributed over the surface. (2) Interactions due to dipole and higher moments may be neglected; also, interactions due to van der Waals' forces are negligible. (3) Electrical interaction ($e\psi$) is small compared with the thermal energy of the ions in the ion atmosphere (kT). The assumption of uniform distribution of charge can only be justified by experimental verification of the theory. The remaining two assumptions are probably nearly valid for protein molecules under ordinary conditions.

Having demonstrated that the mobility under certain conditions is proportional to ϵ , the number of equivalents of H^+ or OH^- bound per mole of protein, it immediately follows that the charge Q on the protein molecule in such instances is given by

$$Q = e\epsilon$$

where e is the electronic charge. Since all quantities in equation 4 are known except r , it is possible to combine mobility and titration data and to calculate by means of equation 4 the "effective" radius, r , of the protein molecule (6, 32, 34).

This calculation has been made for three proteins,—egg albumin, serum albumin, and pseudoglobulin. Results are shown in table 1. Notice that the values calculated (third column) are in fair agreement with those calculated from the molecular volumes (fourth column), and that they are all larger than those obtained from the molecular volume.

Another way of comparing the experimental results with those predicted by the theory is to calculate the theoretical mobility of a sphere of given molecular volume and compare it with the actual mobility. The details of the calculation are omitted, since the equations involved follow directly from those already given. In the fifth column of table 1 is given

⁴ This equation differs from the one given by Henry (20), but follows directly from his method of derivation and corresponds to the function as plotted in his figure. It is concluded that the discrepancy between the printed equation and the one plotted in figure 1 was caused by a typographical error.

the ratio of the experimental to the theoretical mobility for the three proteins considered. It may be seen from the table that, while the experimental value of the mobility is less than the theoretical in all three cases, the agreement is good considering the type of model employed. It is now known that none of these proteins is spherical. However, egg albumin is probably more nearly so than the others, and the agreement, correspondingly, is best for this protein.

Finite size of the ions in the ion atmosphere and the hydration of the protein

In the derivation of equation 2 the ions in the ion atmosphere are implicitly assumed to be point charges. The effect of this assumption and of the hydration of the protein was estimated by one of us (M. H. G.). A short account of the investigation follows

TABLE 1

Radius, in m μ , calculated from titration and mobility data, at 20°C., using the Debye-Hückel theory for a sphere*

| PROTEIN | MOLECULAR WEIGHT† | r _{calcd.} | r _{M.V.} | (V _e /V _t) ₍₂₎ | (V _e /V _t) _(2') |
|---------------------|-------------------|---------------------|-------------------|--------------------------------------------------|---------------------------------------------------|
| Egg albumin..... | 42,200 | 3.26 | 2.82 | 0.592 | 0.615 |
| Serum albumin..... | 67,000 | 3.91 | 2.70 | 0.518 | 0.465 |
| Pseudoglobulin..... | 165,000 | 5.22 | 3.66 | 0.524 | 0.48 |

* Acetate buffers at constant ionic strength (0.02 M for egg albumin and 0.1 M for the other two proteins) were employed. η in equation 4 was corrected for buffer (1.055 η_0 for 0.1 M buffer and 1.011 η_0 for 0.02 M buffer).

† Molecular weights are taken from Svedberg (44).

The ζ -potential, as defined by equation 2, is the potential at the point of closest approach of the ions in the double layer. This point is not on the surface of the protein molecule, but on the surface of an imaginary sphere of radius ($r_p + r_i$), where r_p is the radius of the protein and r_i is an "average" radius of the ions in the double layer. In general, r_i should be expected to be somewhat larger than $r_{M.V.}$, the radius of the protein calculated from the molecular volume due to hydration of the protein. In the development that follows it will be assumed that the protein molecule in solution is surrounded by a monolayer of firmly attached water molecules, through which the ions in the ion atmosphere do not penetrate, and the outside of which forms the surface of slippage between the molecule and the solution. On this basis, then, $r_i = r_{M.V.} + 0.15$, where 0.15 m μ is the diameter of a water molecule. It is obvious, therefore, that the potential, as defined by equation 2, is not the potential required by equation 3, since it is not the potential at the point of slippage.

It is easily shown from the equations of Debye and Hückel that

$$\zeta = \frac{Q(1 + \kappa r_i)}{Dr_p[\kappa(r_p + r_i) + 1]} \quad (2')$$

Equation 2' differs most significantly from equation 2 in the numerator term, $(1 + \kappa r_i)$, which is independent of the size of the protein and is of greatest importance at high ionic strength.

It is of interest to compare the results given by equation 2' with those given by the empirical equation used by Abramson (7) and Moyer and Abels (32) for the determination of an equivalent radius by combining mobility data at two values of the ionic strength. Knowledge of the molecular weight is not necessary for this method. It can easily be shown that equation 2' gives for v_1/v_2 , the ratio of the mobilities at the ionic strengths corresponding to κ equals κ_1 and κ_2 , respectively,

$$\frac{v_1}{v_2} = b \quad \frac{(\kappa_2 r + \kappa_1 r_i + 1)}{(\kappa_1 r + \kappa_1 r_i + 1)} \frac{(1 + r_i \kappa_1)}{(1 + r_i \kappa_2)} \quad (2'')$$

TABLE 2
Comparison of the theoretical and empirical equations

| PROTEIN | RADIUS, IN $m\mu$, CALCULATED FROM | | |
|----------------------|-------------------------------------|-------------|-------------------------------|
| | Equation 2'' | | Empirical equation $a = 3$ |
| | $r_i = 0.3$ | $r_i = 0.3$ | |
| Egg albumin..... | $m\mu$ | $m\mu$ | $m\mu$ |
| 1.6 | 2.0 | 2.2 | |
| Serum albumin..... | 2.7 | 3.4 | 3.4 |
| R-phycoerythrin..... | 3.6 | 6.4 | 5.2 |

where r is the radius of the protein and r_i , as before, is the "average" radius of ions in the ion atmosphere. If v_1/v_2 is known, r can easily be calculated by means of equation 2''. In table 2 equation 2'' is compared with the empirical equation of Abramson (7) and Moyer and Abels (32)

$$\frac{v_1}{v_2} = \left(\frac{\kappa_2 r + a}{\kappa_1 r + a} \right)$$

where a is an empirical factor, the best value of which is 2 (34). Data were taken from Moyer and Abramson (34, table 1 on page 400). It may be seen from table 2 that for $r_i = 0.3$, the values given by equation 2'' and those given by that of Moyer and Abels correspond very closely. The value, 3 Å., for an average radius of the ions in solutions of ordinary electrolytes is a reasonable one. Therefore it may be concluded

that the empirical equation of Moyer and Abels may be represented by the theoretical one (equation 2").

Equation 2' does not improve the agreement between the experimental and theoretical mobility, v_s/v_i . This fact is demonstrated in the last column of table 1, which is headed $(v_s/v_i)_{(2')}$, and is calculated by means of equation 2'. For these calculations r_i was taken as $0.2 \text{ m}\mu$, since they had been completed before it was realized that $0.3 \text{ m}\mu$ was a better value (see table 2).

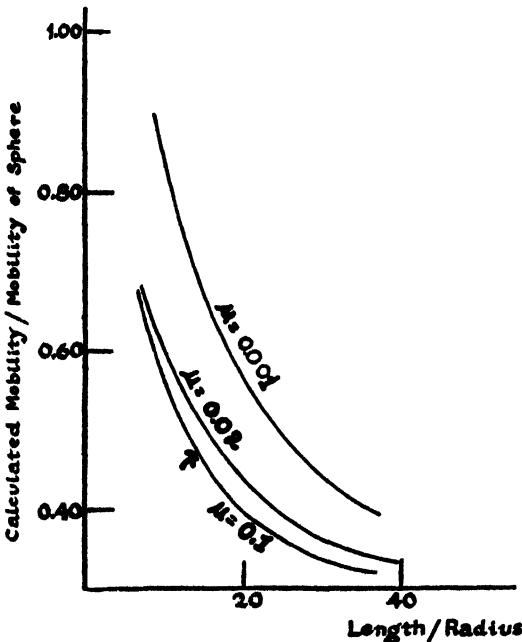


FIG. 11. Calculated mobility of a cylinder having the volume of a serum albumin molecule divided by the calculated mobility of a sphere of the same charge and volume (of the molecule) as a function of the elongation of the cylinder (length divided by the radius) for three values of the ionic strength ($\mu = 0.001, 0.02$, and 0.10).

Calculation of the mobilities of long cylinders

Preliminary estimations of the electrokinetic potential of an aspherical protein molecule were made by using a highly simplified model. Since incomplete data are available for comparison with the theory, this work is offered as the first part of an investigation by one of us (M. H. G.) which may yield independent information about the shapes of protein molecules. Enough progress has been made to indicate that the discrepancy between theory and experiment in the electric mobilities of pro-

tein may be at least partially cleared up by taking their asymmetry into account.

For long cylinders, neglecting the ends, the ξ -potential as a function of the radius and ionic strength is

$$\xi = \frac{2QK_0(ka)}{DlkaK_1(ka)} + \frac{2Q}{Dl} \ln \left(\frac{a+r_i}{a} \right) \quad (5)$$

where Q is the net charge, D is the dielectric constant of the medium, l is the length of the cylinder, and a is its radius. The functions K_0 and K_1 are special solutions of the modified Bessel's equation,⁵ of zero and first order, respectively,

$$\frac{d^2\psi}{dr^2} + \frac{1}{r} \frac{d\psi}{dr} - \left(1 + \frac{n^2}{r^2} \right) \psi = 0 \quad (6)$$

The Debye-Hückel approximations enumerated in the preceding section were used in this development. In applying equation 5, the quantity $a_M \cdot v_r + 0.15$ was used to take into account the hydration of the protein.

For a cylinder migrating parallel to the field, when ka is very large, the mobility is given by

$$v = \xi D / 4\pi\eta \quad (7)$$

according to Henry (20). Henry, however, did not publish the complete analysis for the mobility as a function of ka . A graphical integration of Henry's equations yielded the following for v as a function of ka :

$$v = \xi D / F\pi\eta$$

where F is given in the table below:

| ka | F | ka | F |
|------|------|------|------|
| 0.00 | 8.00 | 1.40 | 6.87 |
| 0.40 | 7.73 | 2.00 | 6.51 |
| 1.00 | 7.36 | 2.40 | 6.34 |
| | | 3.00 | 6.10 |

The mobility therefore becomes

$$v = \frac{2QK_0(ka)}{F\pi\eta lkaK_1(ka)} \quad (9)$$

Equation 9 was applied to a molecule having the volume of serum albumin. In figure 11 the calculated mobilities, at three values of the ionic strength, of cylinders with various degrees of elongation are com-

⁵ The notation is that used by Whitaker and Watson in *Modern Analysis*, p. 373. Cambridge University Press, London (1927).

pared with those calculated for a sphere of the same volume. It may be seen from the figure that the lower the ionic strength the smaller is the difference between the mobility of a sphere and that of a cylinder. However, there is almost no difference between the curves for $\mu = 0.1$ and 0.02. This fact would seem to explain the success of formulation for a sphere, equation 2", in estimating v_1/v_2 for molecules that are definitely not spherical. In addition, it is of interest to note that there is a comparatively small difference between the mobility of spheres and that of highly elongated cylinders having the same net charge and molecular volume.

We may use the available data of Moyer and Abramson (34) for serum albumin at $\mu = 0.1$ to illustrate the use of the theory. The results for this protein require that the mobility be about 61.5 per cent (last column, table 1) of that of a sphere of the same molecular volume. This corresponds, as may be seen in figure 11, to a value of $l/r = 14$, or a ratio of

TABLE 3
Asymmetry from electrokinetic data for cylindrical model

| PROTEIN | LENGTH/BREADTH | |
|--------------------|---------------------|---------------------------------------|
| | Electrokinetic data | Perrin's equation and Svedberg's data |
| Egg albumin..... | 4.4 | 2.6 |
| Serum albumin..... | 6.3 | 5.7 |
| "Globulin"..... | 7.3 | 7.6 |

length to breadth of about 7 for this molecule. When corrected for hydration, the ratio of length to breadth comes out to be 6.2 to 1. Svedberg (45) gives 1.29 for the dissymmetry constant of serum albumin. This, when substituted into Perrin's equation (see Neurath (36)), gives 5.7 for the ratio of length to breadth. The agreement is surprisingly good, considering the simplicity of the model employed. Equation 9 was also applied to the results for pseudoglobulin and egg albumin. The results, together with those for serum albumin, are shown in table 3. Agreement again is excellent.

REFERENCES

- (1) ABRAMSON, H. A. J. Gen. Physiol. **12**, 469 (1928).
- (2) ABRAMSON, H. A. J. Gen. Physiol. **13**, 169 (1929).
- (3) ABRAMSON, H. A. J. Gen. Physiol. **13**, 657 (1930).
- (4) ABRAMSON, H. A. J. Phys. Chem. **35**, 289 (1931).
- (5) ABRAMSON, H. A. J. Gen. Physiol. **15**, 575 (1932).
- (6) ABRAMSON, H. A. J. Gen. Physiol. **16**, 1 (1933).
- (7) ABRAMSON, H. A. J. Gen. Physiol. **16**, 593 (1933).

- (8) ABRAMSON, H. A.: *Electrokinetic Phenomena*. The Chemical Catalog Co., Inc., New York (1934).
- (9) ABRAMSON, H. A., AND GROSSMAN, E. B.: *J. Gen. Physiol.* **14**, 563 (1931).
- (10) ABRAMSON, H. A., AND MICHAELIS, L.: *J. Gen. Physiol.* **12**, 587 (1929).
- (11) ABRAMSON, H. A., AND MOYER, L. S.: *Trans. Electrochem. Soc.* **71**, 135 (1937).
- (12) ABRAMSON, H. A., AND MOYER, L. S.: *J. Gen. Physiol.* **21**, 729 (1938).
- (13) ABRAMSON, H. A., MOYER, L. S., AND VOET, A.: *J. Am. Chem. Soc.* **58**, 2362 (1936).
- (14) ASTBURY, W. T.: *Cold Spring Harbor Symposia Quant. Biol.* **2**, 15 (1934).
- (15) BROSSA, A., AND FREUNDLICH, H.: *Z. physik. Chem.* **89**, 306 (1915).
- (16) BULL, H. B.: *J. Phys. Chem.* **39**, 577 (1935).
- (17) DANIEL, J.: *J. Gen. Physiol.* **16**, 457 (1933).
- (18) DEBYE, P., AND HÜCKEL, E.: *Physik. Z.* **24**, 185 (1923).
- (19) FREUNDLICH, H., AND ABRAMSON, H. A.: *Z. physik. Chem.* **133**, 51 (1928).
- (20) HENRY, D. C.: *Proc. Roy. Soc. (London)* **A133**, 106 (1931).
- (21) HITCHCOCK, D. I.: *J. Gen. Physiol.* **6**, 583 (1923).
- (22) LOEB, J.: *J. Gen. Physiol.* **6**, 395 (1923).
- (23) MATTSSON, S.: *Soil Science* **33**, 41 (1932).
- (24) MOYER, L. S.: *J. Bact.* **31**, 531 (1936).
- (25) MOYER, L. S.: *J. Bact.* **32**, 433 (1936).
- (26) MOYER, L. S.: *Botan. Rev.* **3**, 522 (1937).
- (27) MOYER, L. S.: *Trans. Electrochem. Soc.* **73**, 481 (1938).
- (28) MOYER, L. S.: *J. Phys. Chem.* **42**, 71 (1938).
- (29) MOYER, L. S.: *J. Biol. Chem.* **122**, 641 (1938).
- (30) MOYER, L. S.: *J. Phys. Chem.* **42**, 391 (1938).
- (31) MOYER, L. S.: *Cold Spring Harbor Symposia Quant. Biol.* **6**, 228 (1938).
- (32) MOYER, L. S., AND ABELS, J. C.: *J. Biol. Chem.* **121**, 331 (1937).
- (33) MOYER, L. S., AND ABRAMSON, H. A.: *J. Gen. Physiol.* **19**, 727 (1936).
- (34) MOYER, L. S., AND ABRAMSON, H. A.: *J. Biol. Chem.* **123**, 391 (1938).
- (35) MOYER, L. S., AND MOYER, E. Z.: Annual Report of the Biological Laboratory (Long Island Biological Association, Cold Spring Harbor, Long Island, N. Y.), p. 36 (1937).
- (36) NEURATH, H.: *Cold Spring Harbor Symposia Quant. Biol.* **6**, 196 (1938).
- (37) OTTENBERG, R., AND STENBUCK, F. A.: *J. Gen. Physiol.* **9**, 345 (1925).
- (38) PAULI, W., AND FLECKER, L.: *Biochem. Z.* **41**, 470 (1912).
- (39) PRIDEAUX, E. B. R., AND HOWITT, F. O.: *Proc. Roy. Soc. (London)* **A126**, 126 (1929).
- (40) REINDERS, W., AND BENDIEN, W. M.: *Rec. trav. chim.* **47**, 977 (1928).
- (41) SMITH, E. R. B.: *J. Biol. Chem.* **108**, 187 (1935).
- (42) SMITH, E. R. B.: *J. Biol. Chem.* **113**, 473 (1936).
- (43) STEARN, A. E., AND EYRING, H.: *J. Chem. Phys.* **5**, 113 (1937).
- (44) SVEDBERG, T.: *Chem. Rev.* **20**, 81 (1937).
- (45) SVEDBERG, T.: *Ind. Eng. Chem., Anal. Ed.* **10**, 113 (1938).
- (46) TISELIUS, A.: *Nova Acta Regiae Soc. Sci. Upsaliensis* **7**, No. 4 (1930).
- (47) TISELIUS, A.: *Biochem. J.* **31**, 313 (1937).
- (48) TISELIUS, A.: *Biochem. J.* **31**, 1464 (1937).
- (49) WALPOLE, G. S.: *J. Physiol., Proc. Physiol. Soc.* (1913).
- (50) WINTERSTEINER, O., AND ABRAMSON, H. A.: *J. Biol. Chem.* **99**, 741 (1933).
- (51) WU, H.: *Chinese J. Physiol.* **5**, 321 (1931).

THE WIEN EFFECT: DEVIATIONS OF ELECTROLYTIC SOLUTIONS FROM OHM'S LAW UNDER HIGH FIELD STRENGTHS¹

HARTLEY C. ECKSTROM AND CHRISTOPH SCHMELZER²

Department of Chemistry, Brown University, Providence, Rhode Island

Received September 16, 1938

I. INTRODUCTION

Up to sixteen years ago properties of electrolytic solutions were studied by means of conductance data obtained under low potentials and audio

¹ The symbols used in this article are defined as follows:

$\Lambda_{E=0}$ = ordinary specific conductance at exceedingly low fields,

λ = specific conductance at any field,

$\Lambda_{E=0} = \Lambda_\infty =$ ordinary equivalent conductance at exceedingly low fields at a concentration c ,

Λ = equivalent conductance at any field,

Λ_∞ = equivalent conductance for an infinitely dilute solution at exceedingly low fields,

$\Lambda_{E=\infty}$ = limiting equivalent conductance extrapolated for infinite field,

E = electrical field in E.S.U.,

E' = electrical field in volts per centimeter,

X = x -component of the field in E.S.U.,

V = average potential of the oscillating discharge,

V_F = discharge potential of the spark gap,

$R_{E=0}$ = resistance in ohms at exceedingly low fields,

R = resistance in ohms at any field,

C = capacity,

L = inductance,

D = dielectric constant,

K = dissociation equilibrium constant of the ion pairs,

c' = molar concentration, moles of solute per liter of solution,

c = equivalent concentration, equivalent weight of solute per liter of solution,

m = molal concentration, moles of solute per 1000 g. of solvent,

z_i = valence of the i ion,

Θ = time of relaxation of the ion atmosphere,

θ = decrement (angle),

φ = angle,

κ = characteristic quantity in the Debye-Hückel theory,

ψ = electrical potential around an ion,

$\rho_i = 1/\omega_i$ = frictional resistance of the i ion,

v = frequency,

frequencies. It was never suspected that electrolytes would exhibit interesting and informative properties under other conditions, and it was not until 1927 that Max Wien discovered that the conductivity of an electrolytic solution increases with increasing field strength when very high electrical fields are applied; in other words, that under such extreme conditions Ohm's law is no longer valid.³

This capital discovery, which is of primary importance for the further development of the theory of electrolytic solutions, was the final result of an investigation originally undertaken for the purpose of determining the validity of Stokes' law at high ion velocities. Lenard (16) showed on theoretical grounds that the mobility of an ion should decrease slightly at high speeds. The corresponding decrease of conductivity to be expected for the hydrogen ion, according to theory, was of the order of 0.5 per cent in a field of 1 megavolt per centimeter (10^6 volts per centimeter). The velocity of the hydrogen ion in this field is about 30 meters per second, as compared to about 12 cm. per hour in a field of 1 volt per centimeter. The experimental proof of this prediction presented many difficulties, mainly owing to electrolysis and to the tremendous amount of heat developed in the solution during the time that the resistance is measured. The single-spark method of Wien made such measurements possible; with it Wien showed in 1924 that the conductivity of aqueous solutions of sodium chloride, potassium hydrogen sulfate, and sulfuric acid is independent of field strength within an experimental accuracy of 1 per cent for fields up to 0.5 megavolt per centimeter. This result thus furnished qualitative proof of Lenard's prediction. Later Wien and Malsch, in extending Wien's work to a larger variety of salts and solvents, obtained results that indicated that the temperature coefficient of conductance was somehow influenced by the length of the time interval during which the field was applied (19).

These peculiar results led to a more detailed and accurate study of the phenomenon in question. First, Wien and Malsch developed the barettet method as a new tool which enabled them to measure resistance changes with an accuracy of 1 or 2 parts in 10,000 with fields up to 0.5 megavolt per centimeter (see part II A). Using this method, Wien showed in 1927 that *with high fields the conductance of an electrolytic solution always increases*, and that this effect is not, as was previously supposed, due to an

τ = impulse duration,

ω = frequency in 2π sec.,

η_0 = viscosity of the solvent,

α = degree of dissociation (a parameter in Wilson's theory).

³ University of Jena. Anthony Fellow in Chemistry at Brown University, 1937-38.

* This fact has been known for solid bodies for some time (26).

anomalous temperature coefficient of conductance, but is a general property of electrolytic solutions, dependent upon the concentration and valency of the ions and the specific character of solute and solvent. He was further able to show that for extremely high fields the conductance tends toward a constant limiting value which corresponds nearly, but perhaps not exactly, to that of the equivalent conductivity at infinite dilution as ordinarily measured at low fields. This phenomenon is now known as the first or normal Wien effect (see part II B).

In these earlier investigations only strong electrolytes were studied. When the work was extended to weak electrolytes, another phenomenon appeared, in that the observed effect was many times larger than that for strong electrolytes. A similar effect had already been observed by Gyemant in 1928 in solvents of very low dielectric constant ($D \sim 3$). Wien suggested that the large increase in conductance in the case of weak electrolytes is due to an increase in the degree of dissociation. This new effect is referred to as the dissociation field effect or, sometimes, as the second Wien effect (see part II C).

Both of these effects were discovered without the aid of theory. Subsequently, the observed results have been accounted for fairly satisfactorily on the basis of the theory of ion interaction (see part III).

Following the first experimental work in 1922, Wien and his coworkers devoted themselves to studying all characteristics of the field effect. As their work progressed, it became more and more apparent that this phenomenon would be helpful in leading to a better understanding of electrolytic solutions. Since publications relating to the Wien effect have emanated from many different laboratories and have, in the main, been reported in physical journals, which are not readily accessible to chemists, and since no review of these publications has hitherto been attempted, it seemed worth while to undertake a complete review of the literature and to summarize the results in this new field. It is hoped that the present review will prove helpful to chemists in arriving at an understanding of the underlying problems, many of which give promise of important applications and extensions. In writing this review we have carefully considered all reported studies relating to the Wien effect and have attempted to make the bibliography as complete as possible.

II. EXPERIMENTAL RESULTS

A. *Experimental technique*

The particular experimental method that may best be employed in studying the conductance of electrolytic solutions with high fields is determined by the properties of the solution, such as the specific conductance,

the temperature coefficient of conductance, the density, the heat capacity, and the field strength used. The main difficulty in carrying out these measurements is to avoid excessive heating of the solution (and attendant change in resistance) during the time required for the resistance measurement. Ordinary methods of resistance measurement may sometimes be used for poorly conducting systems and at moderate fields, but generally, for good conductors and at high fields, reliable results are obtained only by means of *impulse methods*, where the high potential acts for a very short period of time. Even then, the influence of the heat effect upon the resistance measurement cannot be completely eliminated. For example, there is still a rise in temperature of 0.24°C . when a field of 100,000 volts per centimeter is applied for only 10^{-6} sec. to a solution with a specific conductance of 10^{-4} mhos (and density and heat capacity of unity). If the temperature coefficient of conductance is 2 per cent per degree, the corresponding resistance change is 0.5 per cent. This example may serve to illustrate the difficulties that had to be overcome in measuring the deviations from Ohm's law.

Resistance measurements during very short time intervals may be carried out in two different ways. First, instantaneous values of the resistance may be recorded by means of a cathode ray oscillograph (Rogowski), and second, the average resistance change may be measured with ordinary integrating instruments (Wien). Whereas the former method is more direct, it cannot compare in accuracy with the second, more tedious, integrating method.

Another difficulty that appears in carrying out these measurements is due to the fact that it is impossible to obtain (with the usual methods) a constant potential for short time intervals (in the order of one-millionth of a second). This, obviously, has to be considered only for the integrating impulse method, where the exact voltage-time function must be known in order that the average resistance may be correlated with an average potential.

In his single-spark method Wien (32) used damped oscillations obtained by discharging a condenser through a resistance and an inductance. Under these conditions the voltage-time function may easily be calculated, provided the electrical constants of the circuit are known.

Figure 1 shows the diagram of the apparatus used by Wien in his original experiments. By opening S_w , in the primary circuit of the induction coil I, the condensers C are charged to a potential which causes a breakdown across the spark gap S. This starts an oscillating discharge in the circuit S-C-L-R₁ (or S-C-L-R₂). The inductively coupled thermocouple Th, together with the galvanometer G, measures the time-integral of a current which is proportional to the current oscillating in the

main circuit. Most of the energy stored in the condensers C is dissipated in the resistance, and therefore the current integral is proportional to the total amount of stored energy,

$$U_0 = V_p^2 C / 2$$

and inversely proportional to the resistance. Two resistances R_1 and R_2 can thus be compared in two consecutive measurements, when the rest of the circuit remains unchanged. If R_1 and R_2 are pure ohmic resistances, then equal galvanometer deflections indicate equality of R_1 and R_2 .⁴ For measurements with electrolytic solutions R_1 represents the conductance cell, containing the solution, and R_2 a comparison resistor, whose value is independent of field strength.

With the single-spark method Wien was able to measure conductivities of electrolytic solutions in fields up to 500,000 volts per centimeter with

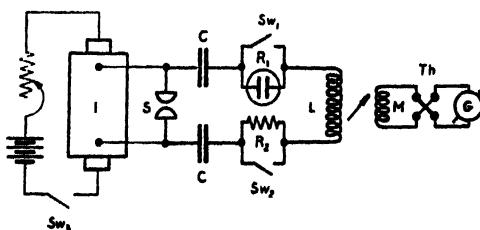


FIG. 1. The single-spark arrangement

an error of about 1 per cent. The accuracy of the method is mainly limited by the relatively poor reproducibility of the spark potential. (Even when radioactive substances are used to ionize the air in the spark gap, it is not possible to eliminate retarded sparks completely.) The field strength of 500 kilovolts per centimeter represents an upper limit determined by the electrical strength of the solution. Moreover, it could only be obtained with special cells (20, 32), for which the calculation of the field strength is very uncertain. In the later quantitative work with ordinary cells, the maximum field strength seldom exceeded 250 kilovolts per centimeter.

In order to enhance precision, a way had to be found that permitted

⁴ The same method can be used for measuring complex resistances (see Malsch: Ann. Physik 84, 841 (1927)). In conductance measurements errors may be introduced by neglecting the inherent capacity of the electrolytic resistance. Its influence increases with increasing frequency, increasing dielectric constant of the solution, decreasing cell constant, and decreasing conductivity. See the discussions of Malsch (17), Malsch and Hartley (18), and Michels (21).

comparison of two resistors simultaneously instead of in two consecutive measurements. Malsch and Wien (20) found the solution of this problem in the baretter method. The fundamental circuit is the same as in the single-spark method, figure 1, but the unknown and the comparison resistor are now connected in parallel. The partial currents through these resistors are compared in a baretter bridge, which reduces the determination of high-frequency currents to a resistance measurement with direct current: the high-frequency current heats a thin wire, thus changing its resistance.

Figure 2 shows a typical baretter arrangement. The main circuit is formed by the spark gap S , inductor L , condenser C , and two parallel arms containing R_1 and R_2 . Each one of these arms is coupled to a baretter B through the resistive coupling unit R_kC_k .⁵ The resistance change of the baretters is measured in the Wheatstone d.c. bridge $B-B-R_h$, which is separated from the main high-frequency circuit by choke coils Ch .⁶

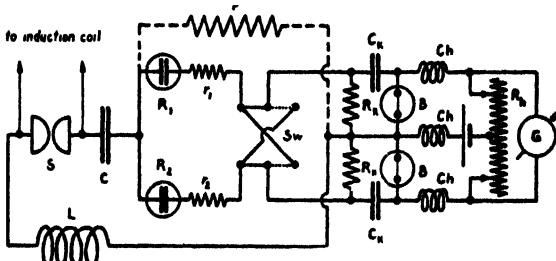


FIG. 2. A typical baretter arrangement

If the resistors in the two arms are equal when the condenser C is discharged, equal currents flow through the two baretters, thus yielding equal changes of resistance, and the d.c. bridge therefore remains balanced. This method is, consequently, a "null" method.

As baretters, Malsch and Wien used ordinary incandescent lamps,⁷ which were selected for having equal resistance as well as equal temperature coefficients. By means of R_h the direct current in the bridge is adjusted to a point where the resistance changes linearly with increasing

⁵ Inductive or capacitive coupling may also be used, but the resistive coupling shown in the diagram has the advantage that one can change the sensitivity over a wide range without largely affecting the electrical constants of the circuit (1).

⁶ Schiele (31) eliminated these chokes by constructing baretters which in themselves constitute small Wheatstone bridges.

⁷ If Wollaston wires of only a few microns in diameter are used in the baretters, the sensitivity can be increased nearly a hundredfold (31).

current, thus allowing linear interpolations between the galvanometer readings. An alternating switch S_w is used to balance out any remaining irregularities.

Types and arrangements of the resistances in both arms are determined by the special problem involved. The combination shown in figure 2 permits the comparison of two electrolytic resistors. The small high-frequency resistors r_1 and r_2 (33) serve to balance the partial currents. If the inherent cell capacity cannot be neglected (see footnote 4), r_1 and r_2 have to be connected parallel to R_1 and R_2 rather than in series, because otherwise large errors may be introduced (18).

The advantage of this baretter method is that the measurements are independent of irregularities of the spark, so that the much higher sensitivity may be fully utilized. An accuracy of 1 to 2 parts in 10,000 as

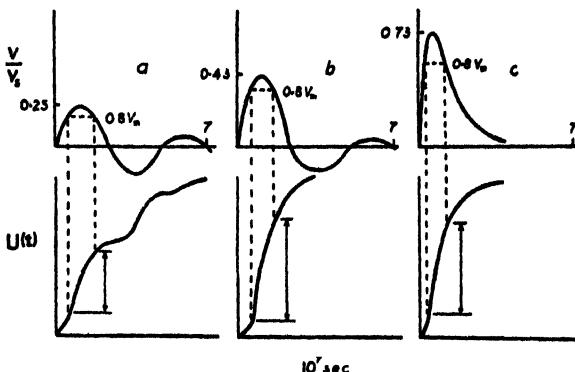


FIG. 3. V/V_s is the ratio of the instantaneous voltage drop across the resistor to the spark potential. $\theta = \pi R \sqrt{C/L}$. In a, $\theta = 1.0$; in b, $\theta = 2.14$; in c, $\theta = 6.28$.

obtained by Schiele (28) probably does not represent the upper limit. Furthermore, this method constitutes a general precision method for measurements with high frequencies and high potentials. It is not confined to resistance measurements alone, but may be used to measure inductances and capacities as well.

For both the single-spark and the baretter method, the form of the voltage-time function is of great importance, because it influences the measured resistance average. The curve should have a flat top, which means that the potential is near its maximum value during a large fraction of the time interval during which the discharge occurs; in other words, most of the energy should be dissipated when the voltage is high. Obviously, this condition has to be fulfilled with any integrating method, and by proper choice of the electrical constants in the oscillating circuit it may be fulfilled. In figure 3, at the top, are shown three typical

voltage-time functions, in which the influence of L , C , and R on their shape is expressed in terms of the logarithmic decrement

$$\theta = \pi R \sqrt{C/L}$$

The lower curves represent the corresponding current integrals or energy-time functions:

$$U(t) = R \int_0^t i^2(t) dt$$

which, for $t = \infty$, are equal to $U_0 = V_p^2 C / 2$. In the case of curve a, only a small amount of the total energy is being dissipated while the voltage drop across the resistor is greater than 0.8 times the maximum voltage V_m (indicated in the figure by vertical arrows). From curve b, however, it is seen not only that V_m increases with increasing decrement, but also that a greater portion of the energy is consumed during the time interval that the voltage is within the desired interval. For still larger decrements V_m increases further, but the amount of energy dissipated in the desired voltage interval now decreases (curve c). For optimum operating conditions it is thus desirable to work with a decrement approximately equal to π .

The impulse duration τ is defined as the time of one half-period of the undamped oscillation of the main circuit.⁸ For this case, $\theta = R = 0$, and

$$\tau = \pi \sqrt{LC}$$

It is seen from figure 3 that for decrements of the order π most of the energy actually is dissipated in such a time interval. With given values for decrement, impulse duration, and total resistance,⁹ the electrical constants that yield the desired voltage-time function are,

$$L = \tau R / \theta \quad \text{and} \quad C = \tau \theta / \pi^2 R$$

The experimentally determined values for the relative conductance change, $\Delta\lambda/\lambda_{s=0}$, at different field strengths represent the actual conductance changes due to deviations from Ohm's law superimposed on con-

⁸ The impulse deviation so defined is only an approximation, since the frequency of a freely oscillating, damped circuit is slightly lower than without damping. The approximation is sufficient for experimental purposes, where only the order of magnitude is of interest. However, it should be considered by comparing experimental data with theoretical results. In this comparison a Fourier analysis of the impulse will indicate whether a correlation of the experimental result with the fundamental frequency, $\omega = \pi/\tau$, is justified or whether the influence of harmonics, $n\omega$, has to be taken into account.

⁹ In the case of the barettter method the parallel resistance of the two arms has to be considered.

ductance changes due to the heating of the solution during the measurement. In order to separate these two effects, the heat effect must be calculated (20) from the energy, U_d , dissipated in the electrolytic resistor. Because of energy losses in the spark itself, U_d is smaller than the total amount of energy stored in the condenser, U_0 ,¹⁰ but a corresponding correction, $p = U_d/U_0$, can easily be determined calorimetrically. Then, the relative change in conductance due to Joule's heat is simply given by

$$\Delta\lambda/\lambda = \epsilon\beta p V_r^2 C / 4m'c^*A$$

where m' is the mass of the solution between the electrodes of the cell, c^* the heat capacity of the solution, and A the mechanical equivalent of heat. The temperature coefficient β is assumed to be independent of field strength. ϵ represents the fraction of energy dissipated in the electrolytic resistor; it is equal to 1/2 in the case of the barettter method, where the energy U_d is dissipated in two equal resistors.

The heat effect may be decreased by decreasing the impulse duration and the energy dissipated in the electrolytic resistor. The latter may be accomplished by inserting a resistor R_p (figure 2) in parallel with the resistors to be measured (21). Both methods are limited because the deviations from Ohm's law depend, in some degree, upon impulse duration, and because the introduction of R_p tends to decrease the sensitivity and influence the decrement of the oscillating circuit. The influence of the heat effect may be eliminated by comparing resistors that have the same temperature coefficient of conductance.

The actual deviations from Ohm's law, as obtained from experiments at high fields after correcting for the heat effect, may be referred to corresponding values of the spark potential V_r . Thus, intercomparisons are possible for measurements that have been carried out with the same decrement and impulse duration. In order to determine the average field strength, which corresponds to a certain measured average conductance, the average potential drop across the resistor must be calculated. This average potential follows from the solution of the fundamental differential equation for a freely oscillating damped circuit, in which the resistance is now a function (assumed known) of the potential. Malsch and Wien (20) considered a linear dependency of conductance and field strength a sufficiently close approximation. They obtained for the average potential in the case of an oscillating discharge ($\theta < 2\pi$)

$$\bar{V} = [4\delta(1 + e^{-\sigma}) / \{3(1 + 2\delta^2)(1 - e^{-\sigma})\}]pV_r$$

where $\delta = \theta/\pi$ and $\sigma = 3\pi\delta(4 - \delta^2)^{-1/2}$. Dividing \bar{V} by the distance between the electrodes in the conductance cell finally yields the average

¹⁰ The losses in the condenser and in the inductance coil are negligible.

field strength. In this computation Malsch and Wien neglect the influence of the stray field around the electrodes. They estimate that uncertainties due to this and the preceding assumption as to the relation between resistance and potential are of the order of several per cent (33).

B. The first Wien effect

There are two types of experiments that show clearly that the conductance of an electrolytic solution increases with increasing field strength.

The first experiment was carried out by Wien (33, 34) with the baretter method and is based on the following principle: The conductance change due to heat is always the same for any form of the voltage-time function, provided the dissipated energy remains unaltered. If, therefore, a change in the form of the voltage-time curve, with otherwise constant parameters

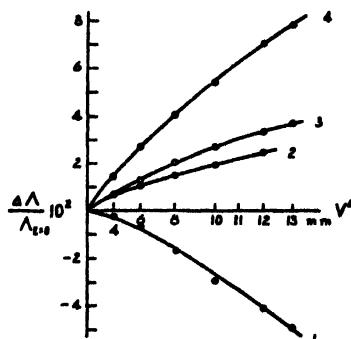


FIG. 4. Acetone solutions of cadmium iodide (curves 1 and 2) and of cobalt chloride (curves 3 and 4), where $\lambda_{E=0} = 0.002$. In all four cases $C = 3029$ cm. In curves 1 and 3 $L = 245,000$ cm., and in curves 2 and 4 $L = 30,000$ cm.

(spark potential and total capacity in the circuit), yields a change of the measured resistance, then this change can only be due to another, superimposed effect which likewise is a function of the field strength.¹¹ Thus, for a positive field effect, the measured resistance change decreases with decreasing decrement (for $0 < \theta < \pi$; see also figure 3). This actually was observed by Wien (33, 34). Figure 4 shows similar results obtained by Bauer (1) for acetone solutions of cadmium iodide (curves 1 and 2) and cobalt chloride (curves 3 and 4) for two values of the decrement. When this decrement was decreased to 0.34 times the initial value (curves 2 and 4), the effect decreased very markedly (curves 1 and 3). In accord-

¹¹ The impulse method is applicable only when relatively long impulses are used, under which condition the impulse duration does not influence the magnitude of the field effect, i.e., for $\tau > 10^{-5}$ sec.

ance with the negative temperature coefficient of conductance of the cadmium iodide solution, the effect even changed its sign.

The experiment by the second method was carried out by Fucks (9) in Rogowski's laboratory, using an instantaneous impulse method. He photographically recorded the voltage drop across an electrolytic resistor with a high-speed cathode ray oscilloscope. Figure 5a shows a diagram of the apparatus used. A condenser is discharged into a long line which terminates at the opposite end in the conductance cell R_1 in series with the constant resistor R_2 . The parallel capacity C balances the effects of the inherent cell capacity and the capacity of the deflection plates in the oscilloscope. For a measurement two consecutive exposures with the same potential are made. In the first, the oscilloscope is connected across BC , which gives the true voltage-time function, whereas in the second, the oscilloscope across AB records the voltage drop across the variable resistor. When the two traces are superimposed, the ratio of the

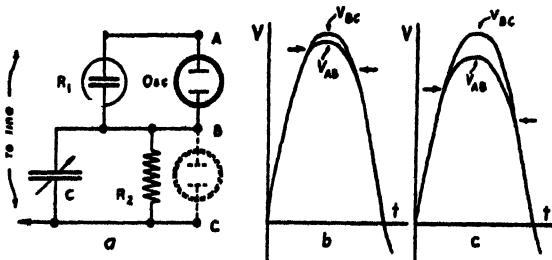


FIG. 5. a, Fucks' oscilloscope arrangement; b and c, drawings of typical traces obtained by Fucks

ordinates measures directly the total resistance change of the solution. Figure 5b shows a result obtained with an aqueous sodium chloride solution, where the field effect is very small compared with the heat effect. The asymmetry of the trace V_{AB} , compared with the trace V_{BC} (indicated by horizontal arrows), shows clearly the heating of the solution during the time interval of 10^{-6} sec.¹² The maximum field strength, E'_m , was 185 kilovolts per centimeter. A similar result for an aqueous solution of barium ferrocyanide, with $E'_m = 180$ kilovolts per centimeter, is shown in figure 5c. Here the deviation from the true voltage-time function, and therefore the total resistance change, is much larger, although the heat effect is nearly the same as in the case of the sodium chloride solution

¹² If R_1 showed only a field effect, the curve V_{AB} would be symmetrical with curve V_{BC} . The observed asymmetry is due to the heat effect, which is an integral function of time.

in figure 5b. This difference can be due only to an actual deviation from Ohm's law. Fucks' results are in excellent agreement with Wien's measurements. This agreement indicates that the method employed by Malsch and Wien to calculate the average potential of an oscillating discharge is correct.¹²

Already the early experiments indicated that the magnitude of the Wien effect increases with increasing valency of the ions. Wien (33) investigated a large number of salts of the valence types 1:1, 1:2, 1:3, 1:4, 2:1, 3:1, 2:2, 2:3, and 3:2 and found as an empirical rule that for the same average field strength and the same conductance at low fields, $\lambda_{E=0}$, the change in conductivity may be expressed by the equation

$$\Delta\Lambda/\Delta_{E=0} = \text{constant } (z_1 z_2)^2 \quad (1)$$

This expression is only an approximation, and deviations as high as 20 per cent may occur for individual salts, especially those of the higher

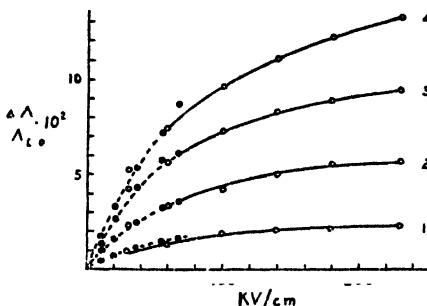


FIG. 6. Curve 1, lithium ferricyanide; curve 2, magnesium sulfate; curve 3, barium ferricyanide; curve 4, barium ferrocyanide. All for $\lambda_{E=0} = 4.5 \times 10^{-5}$. ● and ○ refer to results obtained with two different cells.

valence types. It is, however, very helpful in correcting relative measurements (as obtained with arrangements shown in figure 2) for the Wien effect of the comparison solution. Figure 6 illustrates the influence of valency on the observed conductance change for different salt types.

The dependence of the conductance change on the field strength is somewhat complicated. In figure 7 are shown idealized curves for three different valence types. Initially, at low fields, the curves are convex

¹² Hütter (12) later modified the method of Fucks in such a way that he was able to obtain photographic plots of instantaneous voltage *versus* current. He states that the oscillograph method is limited to an accuracy of about 1 per cent, mainly because it is very difficult to obtain a better concentration of the electronic beam. For more qualitative work, this method gives results much more rapidly than the integrating methods.

towards the axis of potential; they then pass through an inflection point where they approximate linearity and thereafter approach a constant limiting value. In the figure the different portions of the curves are indicated by continuous or broken lines. The limiting characteristic is to be seen also in figure 6, especially in the case of curve 1. How much of the

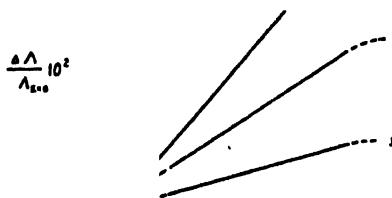


FIG. 7. Curve 1 represents a 2:4 salt type, curve 2 a 2:3 salt, and curve 3 a 1:3 salt

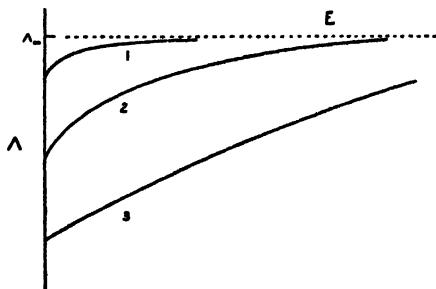


FIG. 8

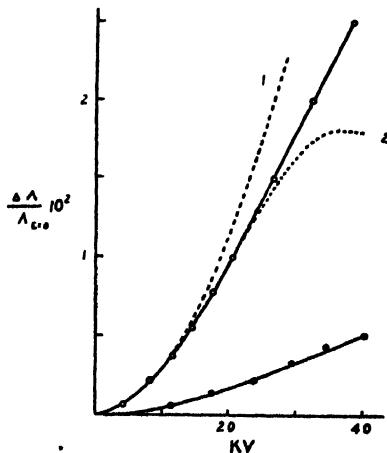


FIG. 9

FIG. 8. Curve 1 represents a very dilute solution, curve 2 a dilute solution, and curve 3 a moderately concentrated solution

FIG. 9. O, barium ferricyanide; ●, lithium ferricyanide

linear and limiting parts of the curves may be obtained experimentally depends upon concentration. In figure 8 the absolute values of conductance are plotted against field strength for three concentrations. For solutions of high concentration, curve 3, the curve is nearly linear throughout, and it is not possible to obtain the limiting effect because of the breakdown of the solution. With solutions of lower concentration, curves

2 and 1, the curves are markedly concave toward the axis of potential; at lower fields the curves rise steeply and at higher fields they flatten out and approach a limiting value. The limit is the same at all concentrations, but this limit is reached at lower fields which are the lower, the lower the concentration.

The limiting effect is the most important result of these investigations. Wien showed that the equivalent conductance of a solution approaches^{1a}

TABLE 1
Limiting field effect and limiting concentration effect of several aqueous solutions

| | | $10^4 c$ | | $\Delta_\infty - \Delta_e$ 10^4 |
|--------------------------|-----|------------|--------------|-----------------------------------|
| $K_4Fe(CN)_6$ | 1:3 | 3.1 6.4 | 3.2 4.3 | 3.0 4.3 |
| $Li_4Fe(CN)_6$ | 1:3 | 3.7 7.5 | 3.7 5.5 | 3.6 4.9 |
| $K_4Fe(CN)_6$ | 1:4 | 2.9 6.4 | 6.1 | 4.7 6.7 |
| $MgSO_4$ | 2:2 | 4.2 8.9 | 6.9 9.0 | 7.5 10.7 |
| $MgCrO_4$ | 2:2 | 4.2 8.7 | 4.4 7.4 | 5.2 7.2 |
| $Ba_2[Fe(CN)_6]_3$ | 2:3 | 3.3 7.2 | 10.3 14.2 | 10.4 15.0 |
| $Ba_2Fe(CN)_6$ | 2:4 | 3.5 7.7 | 12.2 19.8 | 12.9 19.4 |

limiting value as the field is increased indefinitely and that this limiting value is the same as that approached on diluting this solution indefinitely when the conductance is measured at low fields (33, 36, 37) (see, also, Schiele (30)). If Δ_e is the equivalent conductance of an electrolyte at ordinary fields, and Δ_{max} and Δ_∞ are, respectively, its limiting equivalent conductances as the field is increased indefinitely and as the concentration is decreased indefinitely, then $\Delta_{\text{max}} - \Delta_e = \Delta_\infty - \Delta_e$. In table 1 are given values of these differences each divided by Δ_e . As may be seen from

the table, these ratios are equal within the limits of the experimental errors involved in the measurements and the extrapolations. While the limiting field conductance thus agrees with the limiting concentration conductance, it is by no means certain that the two limiting values are identical; while the limiting concentration value is in most cases a close approximation, the limiting field value represents only a rough empirical extrapolation. In this connection attention may be called to Wilson's theory of the Wien effect, according to which a small electrophoretic effect remains at high fields.

In the light of these facts, it follows almost as a necessary conclusion that the high external field tends to decrease those influences which are responsible for the decrease of the ordinary low field conductance with increasing concentration.

The study of the convex part of the curve at low fields (figure 7) was especially difficult, because of the minute changes in conductance that had to be measured. It was therefore possible to study this portion of the curve only for salts of high valence type ($z_1 z_2 > 3$). Wien (35) used a relative method corresponding to that illustrated in figure 2. He corrected for the Wien effect in the comparison solution, which contained a 1:1 salt, in the manner described above, equation 1. The portion of the curve in question has nearly a quadratic form and may be expressed for fields up to 25 kilovolts per centimeter with the first two terms of the following series

$$\Delta\Lambda/\Delta\Lambda_{E=0} = AE'^2(1 - BE'^2 + \dots)$$

where the constants A and B are determined empirically. In figure 9 the experimental curve for barium ferricyanide is shown as a full line.¹⁴ The dotted curves 1 and 2 are computed according to the above formula, using the first and the first and second terms, respectively.

Values of A and B obtained by Wien for aqueous solutions, by Bauer for solutions in acetone, and by Possner for a 50 per cent solution of cane sugar in water as a solvent are given in table 2.

The constants A and B are complicated functions of the temperature, the dielectric constant of the solvent, the valence of the ions, the concentration, the viscosity, and individual ionic properties. Generally, from the limited data available, it appears that A increases and B decreases with decreasing dielectric constant of the solvent. The values of Possner show that viscosity may have a slight influence on the observed results, although this effect is perhaps masked by the influence of the decreased dielectric constant.

¹⁴ The experimental curve for lithium ferricyanide has been inserted to illustrate the influence of the valency of the ions upon the potential effect at low fields.

TABLE 2
Experimental values for A and B

| SALT | VALENCE TYPE | $\lambda_{B=0} \times 10^4$ | $A \times 10^{11}$ | $B \times 10^{11}$ |
|--------------------------------------------------------------------------------------------------------|--------------|-----------------------------|--------------------|--------------------|
| Solvent, water. $T \sim 18^\circ\text{C}$. (Wien (35)) | | | | |
| $\text{Li}_2\text{Fe}(\text{CN})_3$ | 1:3 | 10.0 | 0.51 | 2.1 |
| | | 5.0 | 0.68 | 2.3 |
| | | 2.5 | 0.88 | 2.9 |
| | | 1.25 | 0.96 | 3.8 |
| $\text{Li}_4\text{Fe}(\text{CN})_6$ | 1:4 | 10.0 | 0.91 | 1.5 |
| | | 5.0 | 1.57 | 2.1 |
| | | 2.5 | 2.00 | 2.3 |
| MgSO_4 | 2:2 | 10.0 | 1.11 | 1.7 |
| | | 5.0 | 1.49 | 1.7 |
| | | 2.5 | 1.96 | 2.3 |
| | | 1.25 | 2.50 | 6.0 |
| $\text{Ba}_2[\text{Fe}(\text{CN})_6]_2$ | 2:3 | 10.0 | 2.6 | 1.9 |
| | | 5.0 | 3.9 | 2.1 |
| | | 2.5 | 5.1 | 4.5 |
| | | 1.25 | 5.6 | 7.5 |
| $\text{Ba}_2\text{Fe}(\text{CN})_6$ | 2:4 | 10.0 | 4.2 | 1.9 |
| | | 5.0 | 5.0 | 3.2 |
| | | 2.5 | 10.9 | 3.9 |
| Solvent, acetone. $T \sim 18^\circ\text{C}$. (Bauer (1)) | | | | |
| KI..... | 1:1 | 8.0 | 0.263 | 0.0527 |
| | | 4.0 | 0.334 | 0.0527 |
| | | 2.0 | 0.334 | 0.0438 |
| CdI_3^* | 2:1 | 8.0 | 0.387 | 0.0351 |
| | | 4.0 | 0.615 | 0.0438 |
| | | 2.0 | 0.720 | 0.0351 |
| $\text{Cd}(\text{NO}_3)_2^*$ | 2:1 | 8.0 | 0.211 | 0.0438 |
| | | 4.0 | 0.527 | 0.0438 |
| | | 2.0 | 0.720 | 0.0351 |
| $\text{Ni}(\text{NO}_3)_2^*$ | 2:1 | 8.0 | 0.176 | 0.0527 |
| | | 4.0 | 0.438 | 0.0351 |
| | | 2.0 | 0.527 | 0.0140 |
| Solvent, 50 per cent aqueous solution of cane sugar. $D \sim 60$. 18°C . (Possner (27)) | | | | |
| MgSO_4 | 2:2 | 0.23 | 1.8 | 7.4 |

* These electrolytes are probably not completely dissociated in acetone and thus may exhibit a small dissociation field effect superimposed on the first Wien effect.

While the specific properties of the solution and the temperature influence the magnitude of the Wien effect, it is also dependent upon the length of the time interval during which the field is applied. In table 3 are given values of $\Delta\Lambda/\Lambda_{B=0}$ for three salts for time intervals varying from 6×10^{-7} to 5×10^{-5} sec., the average potential being the same for all time intervals. On examining the table it is seen that in all cases the value of $\Delta\Lambda/\Lambda_{B=0}$ increases initially with increasing length of the time interval and approaches what appears to be a constant limiting value, this limit lying at longer time intervals in the case of salts of higher valence type. Thus for beryllium sulfate $\Delta\Lambda/\Lambda_{B=0}$ has reached a constant value of 1.08 for a time interval of 1.4×10^{-6} sec., while for calcium ferrocyanide a constant value of 10.5 is reached for a time interval of 1.5×10^{-5} sec. This interesting effect has not yet been the subject of a detailed investigation. It has been studied only sufficiently to determine the length of the

TABLE 3

The influence of the impulse duration, τ , on the Wien effect for a constant average potential (33)

| seconds | BeSO_4 | $\text{Ca}_2\text{Fe}(\text{CN})_6$ | $\text{Ba}_2\text{Fe}(\text{CN})_6$ |
|----------------------|-----------------|-------------------------------------|-------------------------------------|
| $\times 10^{-7}$ | 0.92 | | |
| $\times 10^{-7}$ | 0.99 | 4.8 | |
| .4 $\times 10^{-6}$ | 1.08 | 6.8 | 8.6 |
| .5 $\times 10^{-6}$ | 1.04 | 8.4 | 10.9 |
| 1.5 $\times 10^{-6}$ | 1.03 | 10.5 | 13.1 |
| 3.0 $\times 10^{-6}$ | 1.15 | 10.5 | 12.9 |
| 5.0 $\times 10^{-5}$ | 1.07 | 10.7 | 13.1 |

impulse duration necessary to render the observed field effect independent of the length of this impulse duration.¹⁵

Summarizing the outstanding phenomena that characterize the Wien effect, we have: (1) The conductance of a strong electrolyte increases with increasing field strengths and approaches a limiting value which appears to be identical with Λ_∞ . With decreasing concentration this limit is approached at lower fields. (2) The magnitude of the field effect increases with increasing valency of the ions. (3) For very low fields and to a first approximation the conductance increase is proportional to the square of the field, the proportionality constant depending upon the nature of the solute and the solvent. (4) The conductance increase is a function of

¹⁵ This effect led Debye and Falkenhagen (5) to the theoretical prediction of the dispersion phenomena in electrolytic solutions at high frequencies with ordinary potentials.

the impulse duration and for 2:4 salts approaches a constant value for $\tau > 10^{-5}$ sec.; for salts of lower valence type constancy is reached at lower periods.

C. The dissociation field effect

During investigations incident to the development of high electrolytic resistances, Gyemant (11) studied solutions of picric acid in benzene alcohol mixtures. In the course of these investigations he found that the addition of phenol or mineral oils reduces the temperature coefficient of conductance of these solutions to zero. On measuring the resistance of these solutions under high potentials, he observed deviations from Ohm's law which were far greater than any deviations previously observed by Wien and his coworkers. Because of the fact that the temperature coefficient was zero and $\lambda_{E=0}$ was less than 10^{-8} mhos, he determined the resistance by simultaneously measuring voltage and current. Thus, at a field strength of 42 kilovolts per centimeter, he found that for a 0.7 per cent picric acid solution in benzene containing 5 per cent ethyl alcohol and 0.4 per cent mineral oil, $\Delta\Lambda/\Lambda_{E=0}$ was equal to 0.85, and for a 0.7 per cent picric acid solution in benzene containing 8 per cent ethyl alcohol and 3 per cent mineral oil, $\Delta\Lambda/\Lambda_{E=0}$ was equal to 0.52. Deviations of this magnitude have not been observed in aqueous solutions. Furthermore, the conductance of these solutions was found to be a linear function of the field strength.

In order to account for these enormous deviations, Gyemant considered it possible that the degree of dissociation might be increased by high fields. As an *experimentum crucis* he measured the conductance perpendicularly to the high field, reasoning that if the degree of dissociation increases with field strength, the conductivity measured in this direction should increase proportionally to the conductance increase in the direction of the high field. Since the results of this experiment were negative, he concluded that the observed high field effect was probably not due to an increase of dissociation.

Joos (13), relying upon this experiment, suggested that the mechanism might be the same as for the first Wien effect, i.e., a property of free ions only. To clarify this question, Wien (38) and Schiele (28) studied the deviations of strong and weak acids in aqueous solution under the influence of high fields.

These electrolytes were chosen because it had been observed that their solutions exhibit dispersion effects of the same magnitude in high-frequency fields. Both the Debye-Falkenhagen dispersion effect and the first Wien effect are due to interactions between free ions and field. If, therefore, the field effects in strong and weak acids are to be accounted for on the

basis of free ions only, then the magnitude of these effects as observed should be of the same order. If, however, they were found to be of different order, this would indicate that different mechanisms are involved in the effect of fields on strong and weak acids.

Schiele, using an improved baretter method, compared acetic acid and the chloroacetic acids with hydrochloric acid. The dissociation of these acids increases with increasing number of substituent chlorine atoms, and Schiele found that the differences in the conductance change between the acids and hydrochloric acid increase as the carboxylic acid is weaker.

The results of Schiele's investigations are shown in figure 10b, where values of differential conductance effects, $\Delta\Lambda/\Lambda_{E=0}$, are plotted against field strength. Trichloroacetic acid, being a strong acid, shows no observable difference with respect to hydrochloric acid. With dichloroacetic acid (curve 1) a difference is just observable. Monochloroacetic acid

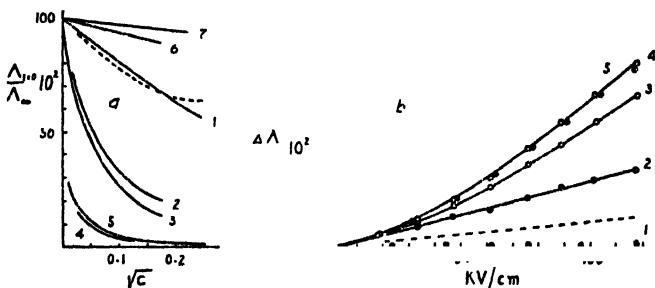


FIG. 10. Curve 1, dichloroacetic acid; curve 2, chloroacetic acid; curve 3, tartaric acid; curve 4, propionic acid; curve 5, acetic acid; curve 6, trichloroacetic acid; curve 7, hydrochloric acid; ----, sulfuric acid. In b, $\lambda_{E=0} = 2.0 \times 10^{-4}$; compared against hydrochloric acid.

(curve 2) shows a large effect and acetic acid (curve 5) a very large one. The curve for propionic acid, which has practically the same strength as acetic acid, lies just below that of the latter acid (curve 4). The curve for tartaric acid lies well below that of acetic acid, which is in accord with the greater strength of tartaric acid. In figure 10a are shown the values of $\Delta\Lambda/\Lambda_0$ for a number of acids. Comparing with figure 10b it is apparent that acids having a low conductance exhibit a high field effect.

The bases exhibit field effects closely paralleling those of the acids. Strong bases exhibit no difference with respect to hydrochloric acid, while weak bases exhibit an effect which is the greater the weaker the base. The effects are illustrated in figure 11.

Solutions of normal salts in solvents of lower dielectric constant in which the electrolyte is incompletely dissociated exhibit the same effect as do

weak acids and bases in water. In figure 12 are shown curves for potassium iodide and lithium bromide in acetone. Potassium iodide is a much stronger electrolyte in this solvent than is lithium bromide, as appears from figure 12a. As may be seen from figure 12b, potassium iodide exhibits a much smaller differential effect than does lithium bromide.

The results just described show conclusively that the magnitude of the field effect increases with decreasing strength of the electrolyte, that is, with increasing association of the ions to ion pairs. Since the large observed increase in conductance cannot be accounted for by the first Wien effect, it is impossible to escape the conclusion that the enhanced effect found in the case of weakly dissociated electrolytes is due to an increased dissociation of the electrolyte under the action of the applied field.

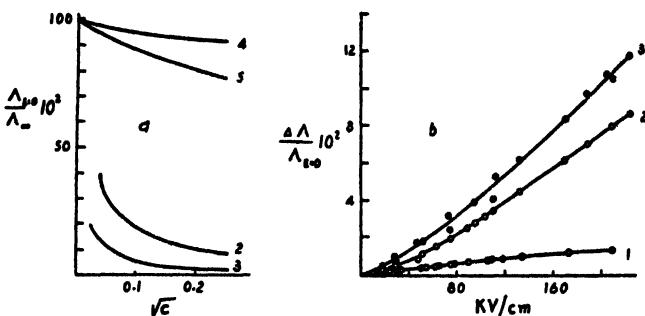


FIG. 11. Curve 1, ●, barium hydroxide, ○, sodium hydroxide; curve 2, methylamine; curve 3, aqueous ammonia; curve 4, sodium hydroxide; curve 5, barium hydroxide. In b, $\Lambda_{E=0} = 2.0 \times 10^{-4}$; compared against hydrochloric acid.

Although the first and the second Wien effects cannot be accounted for by the same mechanism, they, nevertheless, have certain characteristics in common. As in the case of the first Wien effect, the change in conductivity for constant field in the second Wien effect is a function of the impulse duration. Results of Michels (21) for cobalt chloride in acetone are shown in figure 13, from which it is seen that for this electrolyte the effect approaches a constant value for the impulse duration $\tau > 1.7 \times 10^{-6}$ sec.

For low fields, plots of $\Delta\Lambda/\Lambda_{E=0}$ against E' for the dissociation field effect are slightly convex toward the E' -axis, in which respect they resemble the curves of the first Wien effect. On the other hand, no trend toward a limiting effect has been observed in the case of the dissociation field effect. Recalling the $\Lambda_{E=0}/\Lambda_0 - \sqrt{c}$ curves, it is apparent that the dissociation field effect yields only a relatively small increase in the number of free ions at 100 kilovolts per centimeter in the aqueous solutions measured.

Consequently, an extremely high external field would be necessary to dissociate these electrolytes completely;¹⁶ it seems very doubtful that such fields are actually attainable.

In this connection it is of interest to mention the studies of Malsch and Hartley (18) of colloidal solutions of cetylpyridonium chloride in water,

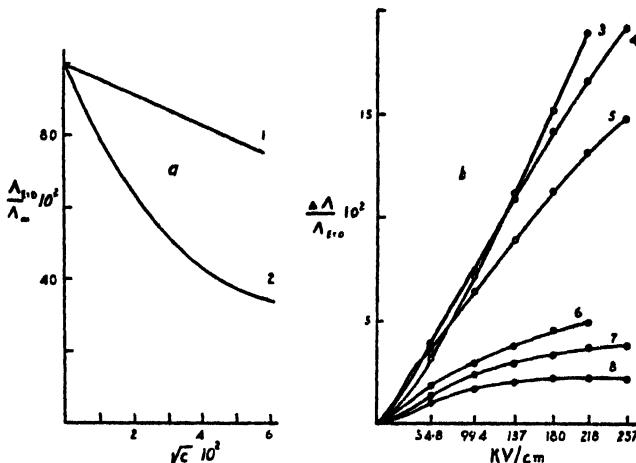


FIG. 12. Acetone solutions of potassium iodide (curves 1, 6, 7, and 8) and of lithium bromide (curves 2, 3, 4, and 5), where in curve 3 $\lambda_{j=0} = 0.00012$ and $c = 0.00168$; in curve 4 $\lambda_{j=0} = 0.00006$ and $c = 0.00058$; in curve 5 $\lambda_{j=0} = 0.00003$ and $c = 0.00023$; in curve 6 $\lambda_{j=0} = 0.00012$ and $c = 0.00073$; in curve 7 $\lambda_{j=0} = 0.00006$ and $c = 0.00036$; and in curve 8 $\lambda_{j=0} = 0.00003$ and $c = 0.00017$.

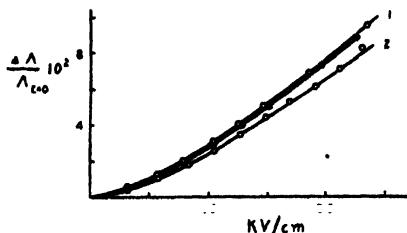


FIG. 13. Acetone solution of cobalt chloride where $\lambda_{j=0} = 0.93 \times 10^{-4}$. In curve 1, O, $\tau = 3.6 \times 10^{-6}$; ●, $\tau = 1.9 \times 10^{-6}$. In curve 2, $\tau = 5.4 \times 10^{-7}$.

where very large potential effects were observed. The outstanding result is that the conductance for very high fields does not approach but exceeds the value of Λ_∞ (figure 14). Malsch and Hartley give the following ex-

¹⁶ It must be remembered that the presence of any free ions leads to an ion atmosphere and thus to a deviation explained by the first Wien effect.

planation: When sufficiently concentrated, the solution contains very large complexes of positive ions (micelles), with some small negative ions relatively firmly attached to these complexes, thus decreasing their mobility and consequently the low field conductance. High fields, however, tend to separate the anions from the large complexes, thus leaving very highly charged positive particles. According to Stokes' law the contribution to the conductance of a complex is larger than that of a corresponding number

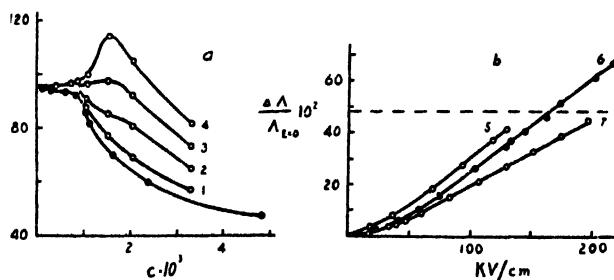


FIG. 14. Cetylpyridonium chloride in water. In a: ●, low potential; curve 1, 50 kilovolts per centimeter; curve 2, 100 kilovolts per centimeter; curve 3, 150 kilovolts per centimeter; curve 4, 200 kilovolts per centimeter; $\tau = 8 \times 10^{-7}$ sec. In b: ----- is $\Delta\Delta/\Delta_{\infty} \cdot 10^2$ for $\Delta = \Delta_\infty; \lambda_{\infty} = 1.29 \times 10^{-4}$; curve 5, $\tau = 16 \times 10^{-7}$ sec.; curve 6, $\tau = 8 \times 10^{-7}$ sec.; curve 7, $\tau = 2 \times 10^{-7}$ sec.

of single ions, so that the conductance may well increase to a value which is larger than Δ_∞ .¹⁷

III. THEORY

Many of the more important properties of electrolytes depend upon ion interaction due to Coulomb forces, which represent the extreme case of a "chemical force" that falls off the most slowly of all forces leading to an interaction. The theoretical analysis of accurate experimental data has

¹⁷ Since the writing of this review, H. Diekmann (Ann. Physik [5]32, 378 (1938)) has published results of an investigation of aqueous solutions of cetylpyridonium chloride and of magnesium sulfate in glycerol-water mixtures at different temperatures. The same effect that is found for cetylpyridonium chloride is also present in systems of magnesium sulfate-glycerol-water. Diekmann therefore draws the conclusion that there is a certain micelle structure in the case of the magnesium sulfate-glycerol-water systems. However, it may be pointed out that: (1) Ordinary conductance data for magnesium sulfate-glycerol-water systems do not show the behavior typical of colloid electrolytes. (2) For the relatively high concentrations studied and for a 2:2 salt in a solvent of $D \sim 60$, relatively strong ionic association must be expected. (3) Diekmann does not mention that the high-field conductance exceeds Δ_∞ , as it does in the case of cetylpyridonium chloride solutions in water.

led to a theory of electrolytes where long-range interactions play a rôle in accounting for the properties of completely dissociated electrolytes, while short-range interactions¹⁸ are an important or determining factor in the case of incompletely dissociated electrolytes. By careful selection of experiments and conditions, various characteristics of these forces become known. Conductance studies, using external fields of low potential and low frequency, of high potential and of high frequency, have been very successful in yielding data which have been theoretically interpreted by using these interactions as the basis of analysis. However, that we know very little about the characteristics of these interactions is apparent when it is recalled that only data for extremely dilute solutions may be accounted for and that even here discrepancies appear.

The Debye-Hückel-Onsager theory of conductance (6, 22) is based on the assumption that only a stationary field of very low potential ($E \sim 0$) is being applied when the conductance is measured. Under these conditions, approximations may be made in which terms containing higher powers of the field are neglected. This leads to an equation in which the conductance does not depend upon the field,—a result which has been amply verified by experiment. Obviously this theory must be modified if it is to account for the first Wien effect, which is observed when the solution is subjected to a high field. As a first attempt in this direction, Joos and Blumentritt (3, 4, 14) evaluated the neglected higher terms of the Debye-Hückel-Onsager equation. They obtained an equation of the form

$$\Delta\Lambda/\Lambda_{E=0} = AX^2 - B'X^4 + \dots$$

This result guided Wien in his investigation of the low-potential region (37). Unfortunately, Joos and Blumentritt made assumptions as a result of which their theory has only a qualitative significance (39). Furthermore, it fails to account for the limiting effect obtained with very high fields.

Using another approach, Falkenhagen (7) later obtained a qualitative theory which applied to the whole curve. His treatment leads to the limiting effect and shows, in general, the influence of valence and dielectric constant. Basing their calculations upon this theory, Falkenhagen and Fleischer (8) introduced the non-stationary field, thus obtaining an equation which shows the dependency of the Wien effect upon frequency.

Overcoming the mathematical difficulties involved, Wilson (39) developed a quantitative theory of the first Wien effect which is applicable to the whole curve; for very low fields the Wilson conductance equation reduces to the quadratic form obtained by Joos and Blumentritt.

¹⁸ Short-range interactions may involve quantum as well as Coulombic forces.

In the case of the dissociation field effect, a different approach must be adopted, since there is no longer complete dissociation, but rather association with an equilibrium existing between ion pairs and free ions. As explained above, the phenomenon of the dissociation field effect was thought to be due to increased dissociation. Onsager (23) showed that while the rate of formation of ion pairs is independent of the field, the rate of dissociation of ion pairs is a function of the field. This analysis indicates that the increased dissociation due to the field is independent of the stoichiometric concentration.

Even though the concepts of ion interaction had been well established in the theory of electrolytes, the extension of the theory to the Wien effect constitutes a major advance in the attack on the general problem of ionic solutions.¹⁹ To review briefly the modern concepts underlying the theory of ionic solutions, we shall first consider the case in which the electrolyte is completely dissociated. The electrostatic attractions and repulsions between the ions do not permit a perfectly random distribution of the ions; on the time average, there will be more ions of unlike sign than of like sign in the neighborhood of the given ion. This "ion atmosphere," as it is called, has a symmetrical distribution when no external field acts. One of the characteristics of this atmosphere is that it has a finite "time of relaxation," i.e., it does not disappear immediately when the central ion is removed, and *vice versa*. Consequently, when a stationary field of very low potential is applied, causing the central ion to move, the distribution of the ion atmosphere becomes asymmetric; there are more ions of like sign in front and unlike sign behind the central ion than in the case of the symmetrical distribution. The "relaxation force" due to asymmetry of the ion atmosphere has a retarding effect upon the motion of the central ion.

There is still another retarding effect acting upon the moving ion, commonly referred to as the "electrophoretic" effect or hydrodynamical effect. According to hydrodynamic laws, moving ions carry solvent with them. This effect extends for some distance from the central ion, so that the ion atmosphere again influences the motion of the central ion. The hydrodynamical effect manifests itself in such a way that the ion is not moving in a stationary solvent, but rather in one that is moving in the opposite direction. These two retarding effects have been incorporated in the Debye-Hückel-Onsager theory of ordinary conductance and account satisfactorily for the decrease of conductance with increasing concentration, in the limiting case of low concentrations.

¹⁹ For a detailed exposition of the theory of electrolytes we refer the reader to the monograph by Falkenhagen (Falkenhagen: Elektrolyte, Hirzel, Leipzig (1932); Electrolytes, Oxford University Press, Oxford (1934); Electrolytes, Alcan, Paris (1934)).

If we apply a very high potential to a solution, we find that the ions are given enormous velocities as a result of which the ion atmosphere is very greatly modified. However, as Wilson shows, in the limit of extremely high fields the ion atmosphere again becomes symmetrical but now contains ions of like sign, ions of unlike sign being constantly swept away at such a rate that the field of the ion in question does not influence the distribution of ions of opposite sign. Thus the relaxation force disappears, leaving only a small contribution from the electrophoretic effect.

Secondly, we shall consider "weak" electrolytes, or "strong" electrolytes in solvents of low dielectric constant. For such systems¹⁰ the energy due to short-range interactions between the ions of unlike sign is great enough, compared with the energy of thermal agitation, so that, on the time average, the two ions spend more time as a pair than as free ions. In other words, the number of ions free to conduct the electric current is small in comparison with the number of ion pairs that have the properties of dipoles. Bjerrum (2) and Fuoss and Kraus (10) have given a satisfactory theory of ionic association which, when combined with the Debye-Hückel-Onsager theory, accounts for the conductance of dilute solutions at low fields. When a high field is applied, two effects come into play: the free ions give rise to a small, normal Wien effect, while the ion pairs undergo increased dissociation in accordance with Onsager's theory of the influence of the field in the dissociation process. In the latter case the high field has shifted the equilibrium, *ion pairs* ⇌ *free ions*, to the right, thus increasing the number of ions free to carry the current. Making a study of this shift in equilibrium, Onsager was able to obtain an equation which is in quantitative agreement with the experimental values.

A. Theory of the first Wien effect

1. The Joos-Blumentritt theory for small fields

Joos and Blumentritt in their first papers (3, 14) calculated the higher terms in the field strengths from the original Debye-Hückel theory (6) and obtained

$$\Delta\Lambda/\Lambda_{E=0} = AX^2 - B'X^4 + \dots$$

where A , B' , ... are known constants. In the later paper by Blumentritt (4) these calculations were repeated, using the modified theory of Onsager (22). The values of these constants are in fair agreement with the experimental values. However, perfect agreement is not possible because Joos and Blumentritt used symmetry conditions which are valid only for zero field. Since it is possible to reduce Wilson's theory to the above form for

¹⁰ We are now considering those systems on the other extreme where $\alpha \ll 1$. There is an intermediate type where α is less than unity, yet is large enough so that the number of free ions is comparable with the number of ions in ion pairs.

low fields, it is not necessary to consider the Joos-Blumentritt theory in detail.

2. The Falkenhagen theory for the complete curve

Under high fields an ion has such a high velocity that it travels through many thicknesses of the ion atmosphere, $1/\kappa$, during a time interval, 2Θ , which is the time necessary for the ion atmosphere to reach a random distribution after the removal of the central ion. Thus, in the case of high fields, the normal ion atmosphere is not formed. It is of interest to calculate the number of ion atmosphere diameters that an ion travels in a high field. This has been done by Falkenhagen (7).

If an i ion with a charge $z_i e = e_i$ moves with a velocity v_i against a frictional resistance ρ_i under a field E , we have

$$v_i = z_i e E / \rho_i = 0.1033 \times 10^{-4} E' l_i / z_i \text{ cm. per second}$$

where E' is the field in volts per centimeter and l_i is the mobility of the ion. If we now multiply v_i by 2Θ , we have the distance the ion travels in 2Θ sec. Furthermore, by dividing this distance by $1/\kappa$, we obtain the distance travelled in terms of the diameter of the ion atmosphere, namely,

$$\kappa s_i = 0.207 \times 10^{-4} E' l_i \Theta \kappa / z_i$$

If we now consider a 10^{-4} molar aqueous solution of potassium chloride, we have, for $18^\circ\text{C}.$,

$$l_1 = l_2 = 65 \text{ ohms}^{-1} \text{ cm.}^2$$

$$\Theta = (0.553/c') \times 10^{-10} \text{ sec.}$$

$$1/\kappa = (3.06/\sqrt{c'}) \times 10^{-8} \text{ cm.}$$

and for $E' = 100$ kilovolts per centimeter we obtain

$$\kappa s_i = 24.4$$

According to this example, then, the time during which a given ion is under the influence of an oppositely charged ion is only one twenty-fifth of the relaxation time, or, in other words, the time during which a given ion is in the neighborhood of other oppositely charged ions is so short that the distribution of the latter remains uninfluenced because of the relaxation effect. The above computation is, of course, only a rough approximation, since it involves assumptions including one such that the relaxation force is proportional to the field, an assumption which does not lead to a limiting effect.

Let us now consider the dependency of the relaxation force upon the external field. We will simplify the problem by neglecting the Onsager

corrections for the influence of the Brownian movement, as well as the electrophoretic effect and its influence upon the relaxation force, and we will further assume that the ions of the electrolyte have equal charge and have the same mobility. Thus the application of this theory is limited as well as approximate.

The equation of continuity yields

$$\frac{dn_i}{dt} = \operatorname{div} \mathbf{v}_i n_i = \operatorname{div} \left(\frac{kT}{\rho_i} \operatorname{grad} n_i - \frac{n_i \mathbf{K}}{\rho_i} \right) \quad (2)$$

where n_i is the time average of the number of i ions per cubic centimeter, and \mathbf{K} is the force acting upon the ions. Let us now change our variables so that the origin of the coördinate system is at the center of the i ion and moves with it, with a velocity v parallel to the X -axis, so that $x = X - vt$, $y = Y$, $z = Z$, $t = t'$,

$$\frac{dn_i}{dt} = -v \frac{\partial n_i}{\partial x} + \frac{\partial n_i}{\partial t'}$$

For the stationary case where $\partial n_i / \partial t' = 0$, and substituting for \mathbf{K} its value $e_i \mathbf{E} - e_i \operatorname{grad} \psi$, equation 2 becomes

$$-v \frac{\partial n_i}{\partial x} = \operatorname{div} \left\{ \frac{kT}{\rho_i} \operatorname{grad} n_i + \frac{n_i e_i}{\rho_i} \operatorname{grad} \psi \right\} \quad (3)$$

The potential ψ is satisfied by the Poisson equation,

$$\Delta \psi = -\frac{4\pi}{D} \sum e_i n_i$$

A further simplification of equation 3 may be effected by introducing for n_i the value $\bar{n}_i + \nu_i$, where \bar{n}_i is the concentration for random distribution and ν_i the perturbation caused by the acting forces. Thus

$$\Delta \nu_i + \frac{\rho_i v}{kT} \frac{\partial \nu_i}{\partial x} = -\frac{n_i e_i}{kT} \Delta \psi \quad (4)$$

where

$$\Delta \psi = -\frac{4\pi}{D} \sum e_i \nu_i$$

For the simple case of an electrolyte having ions of equal charge and equal mobility, we have

$$e_1 = -e_2 = e, \quad n_1 = n_2 = n, \quad \rho_1 = \rho_2 = \rho$$

Letting

$$\nu_1 = -\nu_2 = -\frac{ne}{kT} f, \quad \kappa^2 = \frac{8\pi ne^2}{DkT},$$

and

$$2\mu = \frac{\rho v}{kT} = \frac{eE}{kT}$$

equation 4 becomes

$$\Delta f + 2\mu \frac{\partial f}{\partial x} = \Delta \psi \quad (5)$$

where $\Delta \psi = \kappa^2 f$. Equation 5 may be written in the two forms

$$\Delta f - \kappa^2 f + 2\mu \frac{\partial f}{\partial x} = 0 \quad (6)$$

$$\Delta \left(\Delta \psi - \kappa^2 \psi + 2\mu \frac{\partial \psi}{\partial x} \right) = 0 \quad (7)$$

The distribution function f which satisfies equation 6 and the boundary condition is

$$f = C' \frac{\exp(-\mu x)}{r} \exp\left(-\sqrt{\kappa^2 + \mu^2} r\right)$$

where r is the radial distance from the central ion. To evaluate C' , we let $\mu = 0$ so that

$$(f)_{\mu=0} = C' \frac{\exp(-\kappa r)}{r} \quad (8)$$

When $\mu = 0$, the above equations lead to the original Debye-Hückel expression in which Ohm's law is valid.

In order to obtain C' from equation 8 we recall that for a space charge density P ,

$$\Delta \psi = -\frac{4\pi}{D} P$$

and

$$f = -\frac{4\pi}{D} \frac{P}{\kappa^2} \quad (9)$$

Consequently, for $(f)_{\mu=0}$,

$$(f)_{\mu=0} = -\frac{4\pi P^0}{D \kappa^2} = +\frac{\Delta \psi^0}{\kappa^2} = +\frac{e \exp(-\kappa r)}{Dr}$$

and

$$C' = e/D$$

so that

$$f = \frac{e}{Dr} \exp(-\mu x - \sqrt{\kappa^2 + \mu^2} r) \quad (10)$$

This expression gives the charge density distribution as a function of x and r . What we are now interested in is the force, in the x -direction, due to this distribution. This force is obviously

$$K_x = \frac{e}{D} \int \int \int \frac{P}{r^2} \cos \theta dV$$

With equations 9 and 10 we may reduce the above expression to

$$K_x = -\frac{e^2 \kappa^2}{2D} \int_{-\pi/2}^{\pi/2} \int_0^\infty \frac{\exp(-\mu r \cos \theta - \sqrt{\mu^2 + \kappa^2} r)}{r} \sin \theta \cos \theta d\theta dr \quad (11)$$

and integrating, we obtain

$$K_x = -\frac{e^2 \kappa^2}{D} \left(\frac{w}{2} - \frac{w^2 - 1}{4} \ln \frac{w+1}{w-1} \right) = -\frac{e^2 \kappa^2}{D} y$$

$$\text{where } w^2 = 1 + \frac{\kappa^2}{\mu^2}.$$

In order to study the dependency of y upon μ/κ , we will let $u = (\mu/\sqrt{c'}) \times 10^{-7}$. Plotting y against u , we obtain a curve through the origin which approaches $y = 0.5$ asymptotically as $u \rightarrow \infty$. This means that, for infinite field, $K_x = -e^2 \kappa^2 / 2D$. Since u is proportional to E , we see that a decrease in $\sqrt{c'}$ acts in the same way as an increase in the external field. This is in agreement with the results shown graphically in figure 8.

In order to obtain the conductance in any field, we apply the equation $eE - \rho v + K_x = 0$, so that

$$\lambda = \frac{2ne^2}{\rho} - \frac{2e^2 \kappa^2 n}{\rho D E} y \quad (12)$$

where $2ne^2/\rho = \lambda_\infty$. Since Falkenhagen found it inconvenient to proceed further in an explicit form, he tabulated the functions necessary for calculating the change of conductance with field.

By numerical calculation it may thus be shown that Λ_∞ is approached as a limit at high fields. At lower concentrations this limit is approached at lower fields. Other characteristics of the Wien effect follow from the theory: Increasing the temperature displaces the curves toward higher field strength; a similar shift occurs on decreasing the dielectric constant of the solvent.²¹ The theory accounts for the influence of valency on the field effect, in qualitative agreement with experiment.

²¹ Care should be observed in selecting the system to be studied. Solutions in which ion association occurs should be avoided, since such solutions exhibit the second Wien effect. For example, Bauer's results in acetone seem to be in agreement with the first Wien effect (potassium iodide, for example) and they may well be measurably influenced by ion association. In the case of lithium bromide in acetone the influence of ion association is evident.

As has already been pointed out, all these effects are dependent upon the relaxation time Θ , which decreases with increasing concentration, with increasing temperature, and with decreasing dielectric constant. It is worth while to mention that because of these characteristics of Θ , the Wien effect is closely related to the Debye-Falkenhagen dispersion effect.

3. The Falkenhagen-Fleischer theory for the frequency dependence of the Wien effect²²

Debye and Falkenhagen (5) developed their dispersion theory of electrolytes at low fields by introducing the frequency into the Debye-Hückel-Onsager theory which relates to stationary low fields. In a similar manner, Falkenhagen and Fleischer (8) introduced the frequency into the Falkenhagen theory for stationary high fields as outlined above and thus developed a theory of electrolytes for non-stationary high fields. In so doing, equation 6 has the complex term $\kappa^2(1 + i\omega\Theta)$ instead of just κ^2 . This leads to a differential equation which Falkenhagen and Fleischer have solved, giving

$$\frac{\Delta\Lambda}{\Delta\Lambda_{\mu=0}} = \frac{e^2\kappa^2}{2DkT\mu} y' \quad (13)$$

where the symbols have the same significance as in the preceding theory. y' , which is the real part of a certain function,²³ y^* , is a very complicated expression, which Falkenhagen and Fleischer have expressed in tabular instead of explicit form. Their table contains the real part of y^* as a function of $\mu\sqrt{DT/c'}$ and $\omega\Theta$. In figure 15 are shown curves of y' , the real part of y^* , plotted against $\omega\Theta$ with $\mu\sqrt{DT/c'} \times 10^{-3}$ as a parameter.

Taking the qualitative nature of their theory into account, Falkenhagen and Fleischer conclude that their results are in good agreement with the experimental results of Bauer and Michels. However, the experimental material is so meager that a thoroughgoing comparison between theory and experiment is not now possible. It would seem of interest to study the frequency deviation of the Wien effect under conditions of field and frequency that correspond to the maximum in the $y'-\omega\Theta$ plot.

4. The Wilson theory for binary electrolytes

In formulating his theory Wilson (39) obtained complicated systems of differential equations which he was able to solve and which give a complete

²² Falkenhagen (8) announces a future publication in which he will consider the theoretical treatment of the influence of high fields upon the dielectric constant of the solution. Experimental studies of this phenomenon have not been made as yet and will meet extreme experimental difficulties, at least for solutions of high conductance.

Since the writing of this review, this theory has been published by F. Frölich (*Physik. Z.* **40**, 124 (1939)).

²³ Falkenhagen and Fleischer do not give the function y^* in their publication.

account of the first Wien effect. The solution that Wilson obtained is valid for low-frequency conductance phenomena of binary electrolytes in any field. According to this theory the limiting conductance does not reach Λ_∞ , since only the relaxation force disappears completely, leaving a small contribution due to the electrophoretic effect.

In attacking this problem Wilson proceeded in the familiar way by considering a solution containing per cubic centimeter n_1, n_2, \dots, n_s ions of species 1, 2, ..., s with charges e_1, e_2, \dots, e_s E.S.U., and two elements of volume dV_1 and dV_2 located by a vector \mathbf{r} drawn between them. Considering the time average concentrations of the j ion and i ion in the two elements of volume dV_1 and dV_2 , respectively, it is possible to show²⁴ that

$$f_{ji}(\mathbf{r}) = n_j n_{ij}(\mathbf{r}) = n_i n_{ij}(-\mathbf{r}) = f_{ij}(-\mathbf{r}) \quad (14)$$

where $n_{ij}(\mathbf{r})$ is the time average concentration of the i ions in dV_2 while a j ion is in dV_1 , and $n_{ij}(-\mathbf{r})$ that of the j ions in dV_1 while an i ion is in

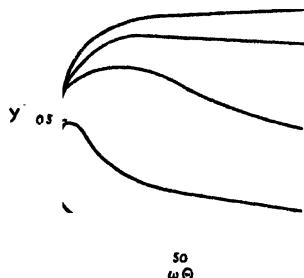


FIG. 15. y' against $\omega\Theta$, where the parameter $\frac{\nu\omega}{2kT} \sqrt{\frac{M_i}{c'}} \times 10^{-9} = 4, 16, 47, 109$, and 1550 (from the bottom up, respectively)

dV_2 . The equations of continuity, in terms of these distributions, are, for the steady state of the solution,

$$\begin{aligned} -\frac{\partial f_{ji}(\mathbf{r})}{\partial t} &= \text{div}_1 f_{ij}(-\mathbf{r}) \mathbf{v}_{ij}(-\mathbf{r}) + \text{div}_2 f_{ji}(\mathbf{r}) \mathbf{v}_{ji}(\mathbf{r}) \\ &= \text{div } f_{ji}(\mathbf{r}) [\mathbf{v}_{ji}(\mathbf{r}) - \mathbf{v}_{ij}(-\mathbf{r})] = 0 \end{aligned} \quad (15)$$

Here $\mathbf{v}_{ji}(\mathbf{r})$ is the total velocity that the i ion has in the vicinity of a j ion and is given by

$\mathbf{v}_{ji}(\mathbf{r}) = \omega_i \mathbf{k}_i - \omega_i kT \text{grad}_2 \ln f_{ji}(\mathbf{r}) - \omega_i e_i \text{grad}_2 \psi_i(0) - \omega_i e_i \text{grad}_2 \psi_i(\mathbf{r})$
where \mathbf{k}_i is the applied force, $\rho_i = 1/\omega_i$ the frictional coefficient, $-\omega_i kT \text{grad}_2 \ln f_{ji}(\mathbf{r})$ the diffusion velocity, $-e_i \text{grad}_2 \psi_i(0)$ the force due to the

²⁴ See references 22 and 25 for details.

ion's own atmosphere, and $-e_i \cdot \text{grad}_i \psi_j(\mathbf{r})$ the force due to the j ion and the latter's atmosphere. There is likewise a similar equation for $\mathbf{v}_{ij}(-\mathbf{r})$. If these equations for $\mathbf{v}_{ij}(-\mathbf{r})$ and $\mathbf{v}_{ji}(\mathbf{r})$ are substituted into equations 15, the following equations of motion are obtained,²⁵

$$\begin{aligned} \text{div } f_{ji}(\mathbf{r})[\mathbf{v}_{ji}(\mathbf{r}) - \mathbf{v}_{ij}(-\mathbf{r})] &= \omega_i[\mathbf{k}_i \cdot \text{grad} f_{ji}(\mathbf{r})] - \omega_j[\mathbf{k}_j \cdot \text{grad} f_{ji}(\mathbf{r})] \\ &- n_i n_j [e_i \omega_i \Delta \psi_j(\mathbf{r}) + e_j \omega_j \Delta \psi_i(-\mathbf{r})] - kT(\omega_i + \omega_j) \Delta f_{ji}(\mathbf{r}) = 0 \end{aligned} \quad (16)$$

If, further, it is assumed that the external field is acting in the direction of the positive x -axis, $\mathbf{k}_1 = X \mathbf{e}_1$, $\mathbf{k}_2 = X \mathbf{e}_2$, and bearing in mind that $e_1 = -e_2 = e$ and $n_1 = n_2 = n$, the following equations are obtained from equation 16:

$$n^2 e \{\Delta[\psi_1(\mathbf{r}) + \psi_1(-\mathbf{r})]\} + 2kT \Delta f_{11}(\mathbf{r}) = 0 \quad (17)$$

$$-n^2 e \{\Delta[\psi_2(\mathbf{r}) + \psi_2(-\mathbf{r})]\} + 2kT \Delta f_{22}(\mathbf{r}) = 0 \quad (18)$$

$$\begin{aligned} Xe(\omega_1 + \omega_2) \frac{\partial f_{12}(\mathbf{r})}{\partial x} - n^2 e \{\omega_1 \Delta \psi_2(\mathbf{r}) - \omega_2 \Delta \psi_1(-\mathbf{r})\} \\ + kT(\omega_1 + \omega_2) \Delta f_{12}(\mathbf{r}) = 0 \end{aligned} \quad (19)$$

$$\begin{aligned} -Xe(\omega_1 + \omega_2) \frac{\partial f_{21}(\mathbf{r})}{\partial x} - n^2 e \{\omega_2 \Delta \psi_1(-\mathbf{r}) - \omega_1 \Delta \psi_2(\mathbf{r})\} \\ + kT(\omega_2 + \omega_1) \Delta f_{21}(\mathbf{r}) = 0 \end{aligned} \quad (20)$$

As boundary conditions: For the flow, the vector field

$$f_{ji}(\mathbf{r})[\mathbf{v}_{ji}(\mathbf{r}) - \mathbf{v}_{ij}(-\mathbf{r})] \quad (21)$$

must be without sources. For the ionic fields, Wilson gives the following:

$$\left. \begin{aligned} \psi_1(\mathbf{r}) - e_1/Dr &< \infty \\ \psi_2(\mathbf{r}) - e_2/Dr &< \infty \end{aligned} \right\} \quad (22)$$

$$\psi_1(\infty) = \psi_2(\infty) = 0 \quad (23)$$

Equations 22 state that, since the space charge of the atmosphere is integrable, the potential $\psi_j(\mathbf{r})$ at the point \mathbf{r} will differ from e_j/Dr only by a finite amount.

The Poisson equations may be expanded into the forms

$$\Delta \psi_1(\mathbf{r}) = \frac{4\pi e}{nD} [-f_{11}(\mathbf{r}) + f_{12}(\mathbf{r})] \quad (24)$$

²⁵ The expansion is accomplished by noting that the \mathbf{k} 's are divergence-free and that the following simplifications have been made. The $\text{grad} \psi(\mathbf{0})$ terms have been neglected and the f_{ji} terms in the $\text{grad} \psi(\mathbf{r})$ terms have been replaced by $n_i n_j$, since $f_{ji} - n_i n_j \sim e_i e_j$. Both of these simplifications are permissible because the terms neglected are of the order e_i^2 , while the other terms are of the order e_i .

$$\Delta\psi_2(r) = \frac{4\pi e}{nD} [-f_{21}(r) + f_{21}(-r)] \quad (25)$$

with similar equations with $(-\mathbf{r})$. In order to obtain one of the two symmetry relations for the potential, one Poisson equation is subtracted from another; this yields

$$\Delta[\psi_1(r) - \psi_1(-r)] = \frac{4\pi e}{nD} [f_{12}(r) - f_{12}(-r)] \quad (26)$$

$$\Delta[\psi_2(r) - \psi_2(-r)] = \frac{4\pi e}{nD} [-f_{21}(r) + f_{21}(-r)] \quad (27)$$

From equation 14 it is seen that

$$\Delta[\psi_2(r) - \psi_2(-r) - \psi_1(r) + \psi_1(-r)] = 0$$

Since $\psi_i(\infty) = \psi_i(-\infty) = 0$, then

$$\psi_2(r) - \psi_2(-r) - \psi_1(r) + \psi_1(-r) = 0$$

or

$$\psi_2(r) - \psi_2(-r) = \psi_1(r) - \psi_1(-r) = 2Y(r) \quad (28)$$

where $Y(r)$ is the odd part of the potential, i.e., the part due to ions of unlike sign.

Since the boundary conditions for the flow stated that the vector field of the flow is without sources, the source-free vector vanishes so that

$$n^2e[\psi_2(r) + \psi_2(-r)] - 2kTf_{22}(r) = \text{Constant}^{28}$$

Since $f_{ii}(\infty) = n^2$, the Constant $= -2kTn^2$. This gives, then,

$$f_{11}(r) = f_{11}(-r) = n^2 - \frac{n^2e}{2kT} [\psi_1(r) + \psi_1(-r)] \quad (29)$$

$$f_{22}(r) = f_{22}(-r) = n^2 + \frac{n^2e}{2kT} [\psi_2(r) + \psi_2(-r)] \quad (30)$$

By means of the Poisson equations and these relations, the following equations are obtained:

$$(\Delta - \kappa^2/2)[\psi_1(r) + \psi_1(-r)] = \frac{4\pi e}{nD} [f_{12}(r) + f_{12}(-r)]$$

$$(\Delta - \kappa^2/2)[\psi_2(r) + \psi_2(-r)] = -\frac{4\pi e}{nD} [f_{21}(r) + f_{21}(-r)]$$

²⁸ Found by writing equation 21 in its expanded forms.

These equations obviously lead to

$$(\Delta - \kappa^2/2)[\psi_1(r) + \psi_1(-r) + \psi_2(r) + \psi_2(-r)] = 0 \quad (31)$$

From the boundary condition (22) it follows that

$$\psi_1(r) + \psi_1(-r) + \psi_2(r) + \psi_2(-r) < \infty$$

As Wilson pointed out, equation 31 has no solution other than zero, which is everywhere finite and free from singularities; thus

$$\psi_1(r) + \psi_1(-r) + \psi_2(r) + \psi_2(-r) = 0$$

or

$$\psi_1(r) + \psi_1(-r) = -[\psi_2(r) + \psi_2(-r)] = 2\Gamma(r) \quad (32)$$

where $\Gamma(r)$ is the even part of the potential. From equations 28 and 32

$$\psi_1(r) = -\psi_2(-r) = \Gamma(r) + Y(r) \quad (33)$$

$$\psi_1(-r) = -\psi_2(r) = \Gamma(r) - Y(r) \quad (34)$$

These equations describe the correct symmetry conditions of the ionic field for a binary electrolyte.²⁷

Denoting the even and odd parts of the function $f_{12}(r)$ by $G(r)$ and $U(r)$, we have

$$f_{12}(r) = f_{21}(-r) = G(r) + U(r) \quad (35)$$

$$f_{12}(-r) = f_{21}(r) = G(r) - U(r) \quad (36)$$

Using these relations and relations 33 and 34, equations 19 and 20 reduce to

$$\frac{n^2 e}{kT} \Delta \Gamma(r) - \Delta G(r) = \frac{eX}{kT} \frac{\partial U(r)}{\partial x} \quad (37)$$

$$\frac{n^2 e}{kT} \Delta Y(r) - \Delta U(r) = \frac{eX}{kT} \frac{\partial G(r)}{\partial x} \quad (38)$$

²⁷ Joos and Blumentritt in their work assumed that the perturbation of the distribution function, f' (where $f = f^0 + f'$) is small compared with f^0 . This assumption is permissible only for low fields, where Ohm's law is valid, and leads to

$$\psi_1(r) + \psi_1(-r) = 0$$

and

$$\psi_1(r) - \psi_1(-r) = 0$$

Blumentritt, by assuming that these expressions are valid at high fields, has limited her results in this way. Consequently her theory gives only the correct sign and trend of the potential effect.

Introducing the symmetry conditions 33 and 34 into 29 and 30, we obtain

$$f_{11}(\mathbf{r}) = f_{22}(\mathbf{r}) = n^2 - \frac{n^2 e}{kT} \Gamma(\mathbf{r}) \quad (39)$$

When equations 33, 34, 39, 35, and 36 are combined with 24 and 25, we have

$$(\Delta - \kappa^2/2) \Gamma(\mathbf{r}) = \frac{4\pi e}{nD} G(\mathbf{r}) \quad (40)$$

$$\Delta Y(\mathbf{r}) = \frac{4\pi e}{nD} U(\mathbf{r}) \quad (41)$$

From equations 38 and 41,

$$(\Delta - \kappa^2/2) U(\mathbf{r}) = -\frac{eX}{kT} \frac{\partial G(\mathbf{r})}{\partial x} \quad (42)$$

and from equations 37 and 40

$$(\Delta - \kappa^2/2) G(\mathbf{r}) - \frac{\kappa^2 n^2 e}{2 kT} \Gamma(\mathbf{r}) = -\frac{eX}{kT} \frac{\partial U(\mathbf{r})}{\partial x} \quad (43)$$

By applying the operator $(\Delta - \kappa^2/2)$ to equation 38 and comparing with equation 42, we finally obtain

$$\Delta(\Delta - \kappa^2) G(\mathbf{r}) = \left(\frac{eX}{kT}\right)^2 \frac{\partial^2 G(\mathbf{r})}{\partial x^2} \quad (44)$$

Thus we have an equation in $G(\mathbf{r})$ alone. Once this $G(\mathbf{r})$ is known, we can obtain $U(\mathbf{r})$, $\Gamma(\mathbf{r})$, and $Y(\mathbf{r})$ from the equations 43, 40, and 41.

With the boundary condition for the flow, equation 21 becomes

$$f_{12}(\mathbf{r})[\mathbf{v}_{12}(\mathbf{r}) - \mathbf{v}_{21}(-\mathbf{r})] = \mathbf{S}_1$$

$$f_{21}(\mathbf{r})[\mathbf{v}_{21}(\mathbf{r}) - \mathbf{v}_{12}(-\mathbf{r})] = \mathbf{S}_2$$

Expanding these and using equations 33, 34, 35, and 36, we obtain for the symmetrical and asymmetrical parts of the flow

$$\mathbf{S}_s = \frac{n^2 e}{kT} \operatorname{grad} \Gamma(\mathbf{r}) - \operatorname{grad} G(\mathbf{r}) - \frac{eX}{kT} \mathbf{e}_s U(\mathbf{r}) \quad (45)$$

$$\mathbf{S}_a = \frac{n^2 e}{kT} \operatorname{grad} Y(\mathbf{r}) - \operatorname{grad} U(\mathbf{r}) - \frac{eX}{kT} \mathbf{e}_s G(\mathbf{r}) \quad (46)$$

where $\mathbf{S}_1 = \mathbf{S}_{s'} + \mathbf{S}_{a'}$, $\mathbf{S}_2 = \mathbf{S}_{s'} - \mathbf{S}_{a'}$, $\mathbf{S}_{s'} = \mathbf{S}_s(\omega_1 + \omega_2)$, $\mathbf{S}_{a'} = \mathbf{S}_a(\omega_1 + \omega_2)$, and \mathbf{e}_s is a unit vector.

Equations 42, 43, 40, 41, 44, 45, 46, together with the ionic boundary conditions

$$\lim_{\alpha \rightarrow 0} \int_{\Omega(S)} \text{grad}_n \Gamma(\mathbf{r}) dS = -\frac{4\pi e}{D}$$

$$\lim_{\alpha \rightarrow 0} \int_{\Omega(S)} \text{grad}_n Y(\mathbf{r}) dS = 0$$

where Ω is the volume enclosed by the surface S , constitute a system of differential equations. On solving these equations we obtain the desired distribution and potential.

Since the field is parallel to the x -axis, cylindrical coördinates (x, ρ, θ) are more convenient to use because the angle variable θ drops out. Then, by Fourier transforms, Wilson finally obtains

$$\begin{aligned} f_{12}(\pm r) &= f_{21}(\mp r) = \frac{2 n^2 \eta e}{\pi D} \left[\int_0^\infty \frac{1}{\sqrt{\kappa^4 - 4\mu^2 \alpha^2}} \right. \\ &\quad \left. \{ (\lambda_1^2 - \alpha^2) K_0(\lambda_1 \rho) - (\lambda_2^2 - \alpha^2) K_0(\lambda_2 \rho) \} \cos(\alpha x) d\alpha \right. \\ &\quad \left. \pm \int_0^\infty \frac{\mu \alpha}{\kappa^4 - 4\mu^2 \alpha^2} \left\{ (\lambda_1^2 - \alpha^2) K_0(\lambda_1 \rho) + (\lambda_2^2 - \alpha^2) K_0(\lambda_2 \rho) \right. \right. \\ &\quad \left. \left. - \frac{\kappa^2}{2} K_0(\lambda_3 \rho) \right\} \sin(\alpha x) d\alpha \right] \end{aligned} \quad (47)$$

$$\begin{aligned} \psi_{1,2}(r) &= \frac{2 e}{\pi D} \left[\int_0^\infty \frac{\kappa^2}{\kappa^4 - 4\mu^2 \alpha^2} \left\{ (\lambda_1^2 - \alpha^2) K_0(\lambda_1 \rho) + (\lambda_2^2 - \alpha^2) K_0(\lambda_2 \rho) \right. \right. \\ &\quad \left. \left. - 4\mu^2 \alpha^2 K_0(\lambda_3 \rho) \right\} \cos(\alpha x) d\alpha \right] \end{aligned} \quad (48)$$

$$\pm \int_0^\infty \frac{\kappa^2 \mu \alpha}{\kappa^4 - 4\mu^2 \alpha^2} \{ K_0(\lambda_1 \rho) + K_0(\lambda_2 \rho) - 2K_0(\lambda_3 \rho) \} \sin(\alpha x) d\alpha \quad (48)$$

where $\eta = e/kT$, $\mu = \eta X$, and $K_0(\lambda \rho)$ is a modified Bessel function of the second kind and zero order, and $\lambda = \lambda(\alpha)$ are functions of α alone.²⁸

$$\lambda_1^2 = \alpha^2 + \frac{\kappa^2}{2} \left(1 + \sqrt{1 - \frac{4\mu^2 \alpha^2}{\kappa^4}} \right)$$

$$\lambda_2^2 = \alpha^2 + \frac{\kappa^2}{2} \left(1 - \sqrt{1 - \frac{4\mu^2 \alpha^2}{\kappa^4}} \right)$$

$$\lambda_3^2 = \alpha^2 + \frac{\kappa^2}{2}$$

$$\lambda_4 = \alpha$$

These expressions will now be used in calculating the electrophoresis and the ionic field.

To obtain the electrophoretic effect,²⁹ the calculation is carried out by using the classical hydrodynamical equation of Stokes for the velocity of a body in a viscous medium,

$$\begin{aligned}\eta_0 \operatorname{curl} \operatorname{curl} \mathbf{v} &= -\operatorname{grad} p + \mathbf{F} \\ \operatorname{div} \mathbf{v} &= 0\end{aligned}\quad (49)$$

where η_0 is the viscosity, \mathbf{v} the vector velocity, p the pressure, and \mathbf{F} the force density due to the action of the external field \mathbf{X} upon the ion atmosphere. The problem is to solve equation 49. With Poisson's equation,

$$F_x = -\frac{DX}{4\pi} \Delta\psi \quad (50)$$

To solve equation 49 we find a vector \mathbf{a} , continuous everywhere, together with its derivatives up to the third order, and satisfying the equation

$$\eta_0 \nabla^2(\nabla^2 \mathbf{a}) = \mathbf{F} \quad (51)$$

Then

$$\mathbf{v} = -\operatorname{grad} \operatorname{div} \mathbf{A} + \nabla \cdot \nabla \mathbf{A} = -\operatorname{curl} \operatorname{curl} \mathbf{a} \quad (52)$$

and

$$p = \nabla^2(\operatorname{div} \mathbf{A}) \quad (53)$$

Using equation 48 in equation 50, together with the above equations, the velocity of the central ion in the direction of the field axis is

$$v_x(0, 0, \theta) = -\frac{Xe\kappa}{6\sqrt{2}\pi\eta_0} f(x) \quad (54)$$

where

$$\begin{aligned}f(x) = 1 + \frac{4}{4\sqrt{2}x^3} \left\{ &-x\sqrt{1+x^2} + \sqrt{2}x \\ &-(1+2x^2)\tan^{-1}\sqrt{2}x + (1+2x^2)\tan^{-1} \\ &+ 2x^2 \sinh^{-1}x \right\} \quad (55)\end{aligned}$$

and $x = \mu/\kappa$. In figure 16a $f(x)$ is shown plotted as a function of x .

This figure shows very clearly how the complicated expression, $f(x)$,

²⁹ See reference 22 and footnote 19 for a discussion.

varies with x , and that $\lim_{x \rightarrow 0} f(x) = \sqrt{2}$ and $\lim_{x \rightarrow \infty} f(x) = 1$. Consequently, if we let $X = 0$, we obtain the result of Onsager (22), whereas, for $X = \infty$,

$$v_s(0, 0, \theta) = -\frac{e\kappa X}{6\sqrt{2}\pi\eta_0}$$

i.e., the electrophoretic effect does not disappear at infinite field.

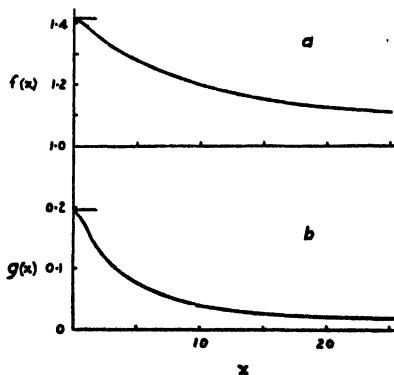


FIG. 16. The functions $f(x)$ and $g(x)$ plotted against x

We obtain the added field acting on the ion in question due to the relaxation force,

$$\delta E(\mathbf{r}) = -\text{grad } \psi_i(\mathbf{r})$$

by using the $\psi_i(\mathbf{r})$ from equation 48. This gives the ionic field at \mathbf{r} , and since we desire the field at the ion itself, we must evaluate the field at $\mathbf{r} = 0$; i.e.,

$$\delta E(0) = e_x \delta X(0, 0, \theta) = \lim_{x \rightarrow \rho=0} \text{grad } \psi_i(\mathbf{r})$$

Performing the indicated operations we have

$$\delta X(0, 0, \theta) = \pm \frac{e\mu\kappa}{2D} g(x) \quad (56)$$

where

$$g(x) = -\frac{1}{2x^3} \left\{ -x\sqrt{1+x^2} + \tan^{-1} \frac{x}{\sqrt{1+x^2}} + \sqrt{2}x - \tan^{-1} \sqrt{2}x \right\} \quad (57)$$

Since $g(x)$ is another complicated expression of x , it is most convenient to show graphically the variation of $g(x)$ with x . This has been done

in figure 16b, from which it is apparent $\lim_{x \rightarrow 0} g(x) = (2 - \sqrt{2})/3$ and $\lim_{x \rightarrow \infty} g(x) = 0$. This simply means that for zero external field the expression 56 reduces to the required finite value and for infinite field the ionic field vanishes.

The final problem is to combine equations 54 and 56 in order to obtain the net velocity of the j ion in a field X .

$$v_j = X \left[e_j \omega_j + \frac{\delta X_j(0)}{X} e_j \omega_j \right] + v_e(0, 0, \theta)$$

Here the first term on the right side is the velocity caused by the external and ionic fields, and $v_e(0, 0, \theta)$ is that caused by the electrophoretic effect.

The mobility of the ion in practical units is

$$u_j = \frac{\omega_j}{300} = \frac{|v_j|}{300X}$$

and since the conductance for infinite dilution $\Lambda_{\infty j} = 96500 u_{\infty j}$, we have as the final equation

$$\Lambda_j = \Lambda_{\infty j} - \frac{|e_j|^2 \kappa}{2DkT} \Lambda_{\infty j} g(x) - (321.67) \frac{|e_j| \kappa}{6\sqrt{2} \pi \eta_0} f(x) \quad (58)$$

In order to show that this equation reduces to the quadratic form for very low fields, we expand the equation

$$\frac{\Delta \Lambda}{\Lambda_{\infty j}} = \left(1 - \frac{|e_j|^2 \kappa}{2DkT} g(x) \right) \frac{\Lambda_{\infty j}}{\Lambda_{\infty j}} - (643.34) \frac{|e_j| \kappa}{6\sqrt{2} \pi \eta_0} \frac{f(x)}{\Lambda_{\infty j}} - 1 \quad (59)$$

for values of $x^2 < 1$. Neglecting terms of x^4 and higher, we have⁵⁰

$$\frac{\Delta \Lambda}{\Lambda_{\infty j}} = \left[(0.0344) \frac{|e_j|^2 \kappa}{2DkT} \frac{\Lambda_{\infty j}}{\Lambda_{\infty j}} + (0.0243)(643.34) \frac{|e_j| \kappa}{6\sqrt{2} \pi \eta_0} \frac{1}{\Lambda_{\infty j}} \right] x^2 \quad (60)$$

We may quantitatively test the Wilson theory by means of Wien's data for magnesium sulfate (35, 36). The $\Lambda_{\infty j} - \sqrt{c}$ plot for magnesium sulfate is linear but does not have the theoretical slope, indicating a slight degree of incomplete dissociation. This manifests itself in such a way that in equation 59 $\lim_{x \rightarrow 0} \frac{\Delta \Lambda}{\Lambda_{\infty j}} \neq 0$. It is therefore necessary to add a constant to equation 59 which has such a value that for $X = 0$ the curves pass through the origin. Equation 59, together with this constant, and equa-

⁵⁰ The Joos-Blumentritt theory gives for binary electrolytes 0.0133 instead of 0.0172 and 0.0973 instead of 0.0248.

tion 60 were used in the calculations. In figure 17 are shown the experimental points as well as the curves computed according to Wilson's theory. In this computation the constants used for the high-field region are $c = 0.37 \times 10^{-3}$, $\Lambda_{E=0} = 105.5$, $\Lambda_\infty = 114$; and in the low-field region $c = 1.22 \times 10^{-3}$, $\Lambda_{E=0} = 99.0$, $\Lambda_\infty = 114$. In all the experimental work in Wien's laboratory the temperature was always about 18°C., unless otherwise specified. In our calculations we therefore assumed $D = 81$ and $\eta_0 = 1.05 \times 10^{-2}$ c.g.s., corresponding to a temperature of 18°C. The deviations at high fields are due to the superimposition of the dissociation field effect, which is insignificant at the lower fields. Unfortunately, accurate data are not available for other electrolytes, and it is therefore impossible to say how generally a similarly good agreement will be found in the case of other electrolytes.

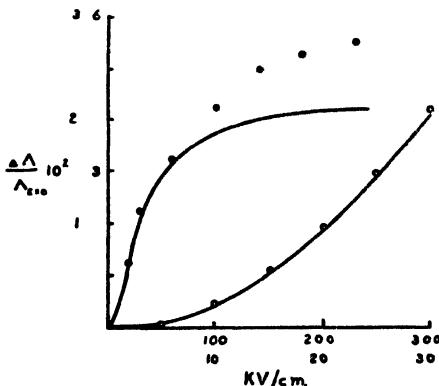


FIG. 17. Aqueous solutions of magnesium sulfate. ●, high field; ○, low field; —, theoretical curves

The theoretical equation (59) is extremely sensitive to the constants. It should therefore be emphasized that, in further studies in this field, accurate values be recorded for concentration, conductance corrected for solvent conductance, and temperature.

B. The theory of the dissociation field effect

In an attempt to account for the results obtained by Wien and his coworkers for weak electrolytes, Onsager (23) made a detailed study of the mechanism of association and dissociation of ions. He assumed that these processes are governed by the laws of Brownian motion. The distinction between free ions and ion pairs is the same as that of Bjerrum (2), who adopted the convention that two ions at a distance $r < q$ are

considered as an ion pair. The potential of the average force between two such ions, i and j , is then

$$w_{ji}(r) = e_j e_i / Dr \quad \text{for } r < q \quad (61)$$

where $q = e_j e_i / 2DkT$. For $r > q$, this potential becomes

$$w_{ji}(r) = \frac{e_j e_i}{D} \frac{e^{(kq - kr)}}{(1 + kq)r} \quad (62)$$

Onsager's considerations apply primarily to cases where the concentration of free ions is sufficiently small, so that

$$\kappa q = -\kappa e_1 e_2 / 2DkT \ll 1 \quad (63)$$

where

$$\kappa^2 = 4\pi(n_1 e_1^2 + n_2 e_2^2) / DkT \quad (64)$$

and n_1 and n_2 denote the concentration of free ions.

When no external field is present, the total concentration of ion pairs equals

$$\nu_{ji} = \nu_{ij} = n_i \int_a^a n_{ji}(r) 4\pi r^2 dr \quad (65)$$

where

$$n_{ji} = n_i e^{-\kappa r/a} \quad (66)$$

is the density of i ions in the neighborhood of j ions, and where a is the distance of closest approach of the two ions.³¹ It is assumed that the ions are rigid spherical point charges, the sum of whose radii is a .

When an external electric field \mathbf{X} is applied, equation 66 is no longer valid. In the presence of such a field no true equilibrium exists between the ions and ion pairs, and it is accordingly necessary to consider the kinetics of dissociation and recombination of ions by applying the equations of Brownian motion (22, 25). The distribution function which results and which is considered to be applicable is the same as that given by equation 14. Again we have for the stationary condition,

$$\operatorname{div}_2 [f_{ji}(\mathbf{r}) \{\mathbf{v}_{ji}(-\mathbf{r}) - \mathbf{v}_{ji}(\mathbf{r})\}] = 0 \quad (67)$$

where $\mathbf{v}_{ji}(\mathbf{r}) - \mathbf{v}_{ji}(-\mathbf{r})$, the mean relative velocity of the two ions at a distance \mathbf{r} , is given by

$$\begin{aligned} \mathbf{v}_{ji}(\mathbf{r}) - \mathbf{v}_{ji}(-\mathbf{r}) &= \omega_i [e_i \mathbf{X} + \mathbf{k}_{ji}(\mathbf{r}) - kT \operatorname{grad}_2 \ln f_{ji}(\mathbf{r})] \\ &\quad - \omega_j [e_j \mathbf{X} + \mathbf{k}_{ij}(-\mathbf{r}) - kT \operatorname{grad}_1 \ln f_{ij}(-\mathbf{r})] \end{aligned} \quad (68)$$

³¹ For a critical discussion of ion pairs and their distribution function, see Fuoss (Trans. Faraday Soc. 30, 967 (1934)).

Here $\mathbf{k}_{ji}(\mathbf{r})$ and $\mathbf{k}_{ji}(-\mathbf{r})$ denote the average forces acting on the ion in question due to interactions with other ions. Assuming equation 63, the screening effect of the ion atmosphere is negligible so that

$$\begin{aligned}\mathbf{k}_{ji}(\mathbf{r}) &= -\text{grad}_2 w_{ji}(\mathbf{r}) \\ \mathbf{k}_{ji}(-\mathbf{r}) &= -\text{grad}_1 w_{ji}(\mathbf{r})\end{aligned}\quad (69)$$

where $w_{ji}(\mathbf{r}) = e_1 e_2 / D r$. Equations 68 and 69 lead to

$$\mathbf{f}(\mathbf{r})\mathbf{v}(\mathbf{r}) = \mathbf{f}(\mathbf{r})[\mathbf{v}_{ji}(\mathbf{r}) - \mathbf{v}_{ji}(-\mathbf{r})]$$

$$= kT(\omega_1 + \omega_2) \left(-\text{grad } f + f \text{ grad} \left(\frac{2q}{r} + 2\beta x \right) \right) \quad (70)$$

where $q = -e_1 e_2 / 2DkT > 0$ and

$$2\beta = |X(e_1 \omega_1 - e_2 \omega_2)| / kT(\omega_1 + \omega_2)$$

The frame of reference is such that the x -axis is parallel to the field \mathbf{X} and $Xe_1 > 0$ and $Xe_2 < 0$. Therefore, equation 67 becomes

$$\text{div grad } f = \text{grad } f \cdot \text{grad} \left(2\beta x + \frac{2q}{r} \right) \quad (71)$$

Furthermore, equation 70 with the equation of continuity yields the net rate of entry of pairs of ions into the interior (Ω) of any closed surface (S) in the \mathbf{r} space, namely,

$$\int_{\Omega} \frac{\partial f}{\partial t} d\Omega = (\omega_1 + \omega_2) kT \int_S \left[\text{grad}_n f - f \text{ grad}_n \left(\frac{2q}{r} + 2\beta x \right) \right] dS \quad (72)$$

where grad_n denotes the normal component of the gradient at S .

If Bjerrum's picture correctly describes the phenomenon of ion association, equation 71 must be solved with the boundary conditions

$$f = n, n_i \quad (r = \infty) \quad (73)$$

and

$$\mathbf{f}(\mathbf{r})\mathbf{v}(\mathbf{r}) = 0 \quad (r = a)$$

whence

$$\text{grad}_r f = f \left(2\beta \cos \theta - \frac{2q}{r^2} \right) \quad \text{for } r = a \quad (74)$$

Since the boundary condition given by equation 74 is complicated, Onsager applied it under the simplifying assumption $r = 0$.

The solution of equation 71 is the subject of another paper by Onsager (24), but instead of solving this equation to satisfy both conditions 73 and 74, he proceeded to calculate separately the parts of f that correspond

individually to the processes of dissociation and recombination. To do this, he obtained the rate constants KA and A in the chemical kinetic equation

$$\frac{\partial \nu_{ii}}{\partial t} = An_i n_i - KA \nu_{ii} \quad (75)$$

which is associated with the reversible reaction $A^+, B^- \rightleftharpoons A^+ + B^-$. The evaluation of these kinetic constants is equivalent to that of Langevin (15). Under the assumptions made, complete dissociation is represented by random distribution, i.e.,

$$f(r) = n_i n_i \quad (76)$$

Accordingly, it follows from equations 72 and 76 that the rate of recombination is

$$An_i n_i = 8\pi qkT(\omega_i + \omega_i) n_i n_i \quad (77)$$

whence for a binary electrolyte $e_1 = -e_2$,

$$A = \frac{4\pi}{D} (e_1^2 \omega_1 + e_2^2 \omega_2) \quad (78)$$

It is now necessary to consider the case where the field X vanishes, in order to obtain that part of the distribution function which represents the associated ions. In connection with equation 62, Onsager points out that the factor e^{-qr} is due to the gradual screening off of the electric field of an ion by the ion atmosphere and that neglect of this factor would lead to an infinite total space charge. Taking into account assumption 63, it is possible to neglect this factor up to some distance, r' , which fulfills the condition

$$a \ll r' \ll 1/\kappa \quad (79)$$

so that

$$w_{ii}(r) = e_i e_i / Dr \quad \text{for } r < r'$$

Now, according to condition 65, 76, and 79, the part of the distribution function representing the associated ions is

$$f(r) = n_i n_i (e^{2q/r} - 1) = \nu_{ii} K(0) (e^{2q/r} - 1) \quad (80)$$

This equation is the only solution of equation 71 that satisfies the following set of boundary conditions: (1) $f = 0$ for $r = \infty$; (2) there is a source of the given yield at the origin; and (3) the flow from the origin through any space angle is finite. Onsager states, however, that the solution of equation 71 for the field $X \neq 0$ under these same boundary conditions

involves elaborate analysis, but that the unique result is expressed by the definite integral

$$f(r, \theta) = \frac{g}{r} e^{-\beta r + \beta r \cos \theta + 2q/r} \int_{s=0}^{s=2q} J_0[(-8\beta s)^{1/2} \cos \frac{1}{2}\theta] e^{-s/r} ds \quad (81)$$

where $J_0[]$ denotes the ordinary Bessel function of order zero. If r is small, the upper limit³² may be replaced by ∞ to yield the result

$$f(r, \theta) \sim g e^{2qr} e^{2\beta r \cos \theta} \quad (82)$$

Onsager further reduces this expression by replacing $e^{2\beta r \cos \theta}$ by unity, stating that the error introduced thereby may be of the same order of magnitude as that caused by neglecting the hydrodynamic interaction of the ions. Comparing the resulting simplified expression with equation 80, where $e^{2q/r} \gg 1$, it follows that

$$\nu_{ii} = g/K(0) \quad (83)$$

It is now necessary to evaluate the rate of dissociation by substituting the distribution 81 into the rate equation (72). If the integral in equation 81 is written $\int_0^{2q} = \int_0^{\infty} - \int_{2q}^{\infty}$, only the integral $-\int_{2q}^{\infty}$ contributes to the flow, as Onsager points out, so that for small r ,

$$-K(X)A\nu_{ii} = -8\pi qkT(\omega_i + \omega_s)g \frac{J_1(4\sqrt{-\beta q})}{2\sqrt{-\beta q}}$$

Using equations 78 and 84, this reduces to

$$\frac{K(X)}{K(0)} = \frac{J_1(4\sqrt{-\beta q})}{2\sqrt{-\beta q}} = F(b) \quad (84)$$

Here

$$F(b) = 1 + b + \frac{b^2}{3} + \frac{b^3}{18} + \frac{b^4}{1800} + \frac{b^5}{2700} + \frac{b^6}{56700} + \dots \quad (85)$$

for small values of b , or

$$F(b) \sim \sqrt{\frac{2}{\pi}} \frac{e^{(8b)^{1/2}}}{(8b)^{-3/4}} \left[1 - \frac{3}{8(8b)^{1/2}} - \frac{15}{128(8b)} - \frac{105}{1024[(8b)^{1/2}]^3} - \dots \right] \quad (86)$$

for large values of b . Equation 84 gives the increase of the dissociation constant due to the field, where the parameter b is

$$b = 2\beta q = \frac{z_1\omega_1 + z_2\omega_2}{\omega_1 + \omega_2} z_1 z_2 \frac{|X| e^3}{2Dk^2 T^2}$$

³² The error thus caused is negligible if

$$\frac{2q - 4(\beta q)^{1/2}}{e^r} >> 1$$

Before proceeding further, it may be well to point out that equation 84 is an approximation which results on (1) assuming $a = 0$ for the boundary condition (74), and (2) obtaining equation 82 by replacing the upper limit in equation 81 by ∞ and setting $e^{\frac{2\beta r}{\alpha} \cos \theta} = 1$ in equation 82.³³ It may also be noted that the earlier equation (71) itself expresses the result of a simplified picture: (1) The hydrodynamic interaction of the ions is neglected, which may cause important errors in any solvent if the field is very high. (2) The shielding of the Coulomb forces due to the ion atmosphere has been neglected.

Physically, equation 84 states that while the dissociation of ion pairs is increased by the field, the recombination of ions remains unaffected (see equation 77). Considering the case of binary electrolytes, it is seen that

$$\frac{c\alpha^2}{1 - \alpha} = K = K_0 F(b) \quad (87)$$

For small values of b , a good approximation is

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta\alpha}{\alpha_0} = \left(\frac{1 - \alpha}{2 - \alpha_0} \right) \frac{\Delta K}{K_0} = \frac{1 - \alpha}{2 - \alpha_0} b \quad (88)$$

where the increase of conductance λ is assumed proportional to the displacement of the dissociation equilibrium.

When the ions have the same charges, and the total number of ions is small compared to that of ion pairs, i.e., $\alpha \ll 1$, then

$$\frac{\lambda}{\lambda_{B=0}} = \frac{\alpha}{\alpha_0} = \sqrt{\frac{K}{K_0}} = \sqrt{F(b)} = 1 + \frac{1}{2}b + \frac{1}{24}b^2 + \dots \quad (89)$$

It is of interest to point out that according to equation 89 the concentration of the electrolyte and the dimensions of the ions, as well as the equilibrium constant, have no influence on the value of $\lambda/\lambda_{B=0}$. In other words, for a given field this ratio depends only upon the charges on the ions, their mobility, the dielectric constant of the medium, and the temperature. Considering the case of a 1:1 salt under the action of the

³³ See footnote 32. If equation 71 is assumed valid for all distances, $r = a$, and if

$$\frac{q^2}{a^2(1 - a\sqrt{\beta/q})} e^{-2q(1-a\sqrt{\beta/q})^{1/2}/a}$$

is small, the boundary condition for small r is immaterial. Since the errors caused by these simplifying assumptions are governed by values of β and q , we must consider the limits of these quantities. Since the intensities of the electrical field do not exceed 500 kilovolts per centimeter, β is limited to $0 < \beta < 10^7 \text{ cm.}^{-1}$ However, since q depends on the valence type of the electrolyte and on the dielectric constant of the solvent, considerable error may appear in equation 84 in solvents of high dielectric constant like water and acetone.

extreme field of 500 kilovolts per centimeter at 25°C., the resulting value obtained is $\lambda/\lambda_{x=0} = 1.37$ for water and 121 for benzene. This may account for the large effect observed by Gyemant.

Onsager compares the consequences of this theory with available experimental data by means of equations 87 and 89, and discusses the results at some length. In figure 18, reproduced from Onsager's paper, are shown curves based on the experimental values of Schiele (28) for acetic and chloroacetic acids in aqueous solution. As may be seen from the figure, there is excellent agreement between the theoretical curve as drawn and the experimental points, except at low fields where deviations appear. This discrepancy is due to various simplifications and assumptions introduced in the course of developing the theory.

Necessary data are lacking for solutions in solvents of quite low dielec-

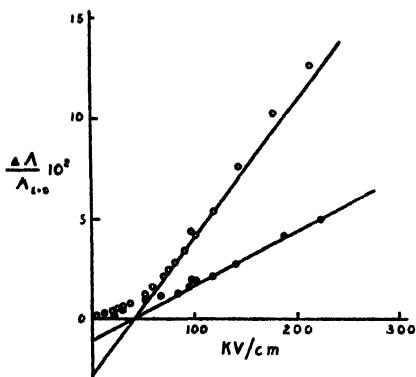


FIG. 18. ○, acetic acid; ●, chloroacetic acid; —, theory

tric constant. Fuoss and Mead³⁴ have measured the conductance of solutions of tetrabutylammonium picrate in diphenyl ether at 50°C. up to potentials of 20 kilovolts per centimeter. With a solution having a concentration of 7.14×10^{-3} moles per liter, they found that the conductance curve closely approximates the theoretical curve. They found, however, contrary to theory,³⁵ that the slope of the curve is a function of concentration and at other concentrations the variation amounts to as much as 40 per cent.

³⁴ These experiments have been carried out in the Research Laboratories of the General Electric Company and are as yet unpublished. We are indebted to Dr. R. M. Fuoss for this information.

³⁵ Onsager recently stated that perhaps by a more detailed analysis his theory could be extended to account for this concentration effect.

IV. CONCLUSION

The deviations from Ohm's law for electrolytic solutions, discovered by Wien and exhaustively investigated by him and his coworkers, may be accounted for by the present theory of Coulombic interaction between the ions. The agreement between experiment and the theory of Wilson is remarkably good, considering the complexity of the hydrodynamical equations involved in the theoretical treatment and the difficulties encountered in the experimental determinations.

Although the present experimental material includes results covering most of the variable parameters involved, further studies of the dependence of the general Wien effect upon viscosity, dielectric constant, and temperature are greatly needed. The influence of frequency in the high-frequency ranges requires further and systematic investigation, particularly in view of the recent theory of Falkenhagen and Fleischer.

It is well established that, in solutions of weak electrolytes, the field effect is much larger than in solutions of strong (completely dissociated) electrolytes. The parallelism found between the field effect and the strength of the electrolyte leaves little room for doubt that the external field increases the number of free ions in the solution. Onsager's theory of the field effect in solutions of weak electrolytes accounts remarkably well for the observed experimental results in the few instances where the necessary data are available for making the comparison. Investigation of the field effect in solutions of incompletely dissociated electrolytes is much needed, particularly for solutions in solvents of low dielectric constant where the phenomenon is greatly simplified owing to the low concentration of free ions.

The investigation of the influence of the electric field on the properties of electrolytic solutions, theoretical as well as experimental, has served not only to disclose many new and interesting phenomena, but also to place the theory of these solutions on a thoroughly secure foundation. The ions as well as the products of their interaction have taken on a degree of physical reality, the lack of which previously made it so difficult for many chemists to accept the ionic theory.

Gratitude is expressed to Professors Charles A. Kraus and Lars Onsager, who have read the manuscript and have offered helpful suggestions.

REFERENCES

- (1) BAUER, F.: Ann. Physik [5] 6, 253 (1930).
- (2) BJERRUM, N.: Kgl. Danske Videnskab. Selskab Math. fys. Medd. 7, 9 (1926).
- (3) BLUMENTRITT, M.: Ann. Physik 85, 812 (1928).
- (4) BLUMENTRITT, M.: Ann. Physik [5] 1, 195 (1929).

- (5) DEBYE, P., AND FALKENHAGEN, H.: *Physik. Z.* **29**, 121, 401 (1928).
- (6) DEBYE, P., AND HÜCKEL, E.: *Physik. Z.* **24**, 305 (1923).
- (7) FALKENHAGEN, H.: *Physik. Z.* **30**, 163 (1929); *Physik. Z.* **32**, 353 (1931).
- (8) FALKENHAGEN, H., FRÖLICH, F., AND FLEISCHER, H.: *Naturwissenschaften* **25**, 446 (1937).
- FALKENHAGEN, H., AND FLEISCHER, H.: *Physik. Z.* **39**, 305 (1938).
- (9) FUCKS, W.: *Ann. Physik* **12**, 306 (1932).
- (10) FUOSS, R. M.: *Chem. Rev.* **17**, 27 (1935).
FUOSS, R. M., AND KRAUS, C. A.: *J. Am. Chem. Soc.* **55**, 476 (1933).
- (11) GYEMANT, A.: *Wiss. Veröffentl. Siemens-Werken* **7**, 134 (1928); *Physik. Z.* **29**, 289 (1928).
- (12) HÜTER, W.: *Ann. Physik* **24**, 253 (1935).
- (13) JOOS, G.: *Physik. Z.* **29**, 570 (1928).
- (14) JOOS, G.: *Physik. Z.* **29**, 755 (1928).
JOOS, G., AND BLUMENTRITT, M.: *Physik. Z.* **28**, 836 (1927).
- (15) LANGEVIN, P.: *Ann. chim. phys.* **28**, 433 (1903).
- (16) LENARD, P.: *Ann. Physik* **41**, 53 (1913); **61**, 665 (1920).
- (17) MALSCH, J.: *Ann. Physik* **84**, 841 (1927); *Physik. Z.* **29**, 770 (1928).
- (18) MALSCH, J., AND HARTLEY, G. S.: *Z. physik. Chem. A170*, 321 (1934).
- (19) MALSCH, J., AND WIEN, M.: *Physik. Z.* **25**, 559 (1924).
- (20) MALSCH, J., AND WIEN, M.: *Ann. Physik* **83**, 305 (1927).
- (21) MICHELS, F.: *Ann. Physik* **22**, 735 (1935).
- (22) ONSAGER, L.: *Physik. Z.* **27**, 388 (1926); **28**, 277 (1927).
- (23) ONSAGER, L.: *J. Chem. Phys.* **2**, 599 (1934).
- (24) ONSAGER, L.: Dissertation, Yale University, 1935.
- (25) ONSAGER, L., AND FUOSS, R. M.: *J. Phys. Chem.* **36**, 2689 (1932).
- (26) POOLE, H. H.: *Phil. Mag.* **32**, 112 (1916); **34**, 204 (1917); **42**, 488 (1921).
- (27) POSSNER, H.: *Ann. Physik* [5] **6**, 875 (1930).
- (28) SCHIELE, J.: *Ann. Physik* **13**, 811 (1932).
- (29) SCHIELE, J.: *Physik. Z.* **34**, 60 (1933).
- (30) SCHIELE, J.: *Physik. Z.* **34**, 61 (1933).
- (31) SCHIELE, J.: *Physik. Z.* **35**, 632 (1934).
- (32) WIEN, M.: *Physik. Z.* **23**, 399 (1922); *Ann. Physik* **73**, 161 (1924).
- (33) WIEN, M.: *Ann. Physik* **83**, 327 (1927).
- (34) WIEN, M.: *Physik. Z.* **28**, 834 (1927).
- (35) WIEN, M.: *Ann. Physik* **85**, 795 (1928).
- (36) WIEN, M.: *Physik. Z.* **29**, 751 (1928).
- (37) WIEN, M.: *Ann. Physik* [5] **1**, 400 (1929).
- (38) WIEN, M.: *Physik. Z.* **32**, 545 (1931).
- (39) WILSON, W. S.: Dissertation, Yale University, 1936.

